02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS

Submit only ONE copy of this form **for each PI/PD and co-PI/PD** identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. ***DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.***

PI/PD Name: Ewa Deelman

Gender: ☐ Male ☒ Female

Ethnicity: (Choose one response) ☐ Hispanic or Latino ☒ Not Hispanic or Latino

Race: (Select one or more)
☐ American Indian or Alaska Native
☐ Asian
☐ Black or African American
☐ Native Hawaiian or Other Pacific Islander
☒ White

Disability Status: (Select one or more)
☐ Hearing Impairment
☐ Visual Impairment
☐ Mobility/Orthopedic Impairment
☐ Other
☒ None

Citizenship: (Choose one) ☒ U.S. Citizen ☐ Permanent Resident ☐ Other non-U.S. Citizen

Check here if you do not wish to provide any or all of the above information (excluding PI/PD name): ☒

REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project ☒

Ethnicity Definition:
Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.
Race Definitions:
American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
Black or African American. A person having origins in any of the black racial groups of Africa.
Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

WHY THIS INFORMATION IS BEING REQUESTED:

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information recieved from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational oppurtunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS

Submit only ONE copy of this form **for each PI/PD and co-PI/PD** identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. ***DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.***

PI/PD Name: Burt Holzman

Gender: ☐ Male ☐ Female

Ethnicity: (Choose one response) ☐ Hispanic or Latino ☐ Not Hispanic or Latino

Race: (Select one or more)
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Black or African American
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ White

Disability Status: (Select one or more)
- ☐ Hearing Impairment
- ☐ Visual Impairment
- ☐ Mobility/Orthopedic Impairment
- ☐ Other
- ☐ None

Citizenship: (Choose one) ☐ U.S. Citizen ☐ Permanent Resident ☐ Other non-U.S. Citizen

Check here if you do not wish to provide any or all of the above information (excluding PI/PD name): ☒

REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project ☐

Ethnicity Definition:
Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

Race Definitions:
American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
Black or African American. A person having origins in any of the black racial groups of Africa.
Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

WHY THIS INFORMATION IS BEING REQUESTED:

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information recieved from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational oppurtunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

02 INFORMATION ABOUT PRINCIPAL INVESTIGATORS/PROJECT DIRECTORS(PI/PD) and co-PRINCIPAL INVESTIGATORS/co-PROJECT DIRECTORS

Submit only ONE copy of this form **for each PI/PD and co-PI/PD** identified on the proposal. The form(s) should be attached to the original proposal as specified in GPG Section II.C.a. Submission of this information is voluntary and is not a precondition of award. This information will not be disclosed to external peer reviewers. ***DO NOT INCLUDE THIS FORM WITH ANY OF THE OTHER COPIES OF YOUR PROPOSAL AS THIS MAY COMPROMISE THE CONFIDENTIALITY OF THE INFORMATION.***

PI/PD Name: Frank Wuerthwein

Gender: ☐ Male ☐ Female

Ethnicity: (Choose one response) ☐ Hispanic or Latino ☐ Not Hispanic or Latino

Race: (Select one or more)
- ☐ American Indian or Alaska Native
- ☐ Asian
- ☐ Black or African American
- ☐ Native Hawaiian or Other Pacific Islander
- ☐ White

Disability Status: (Select one or more)
- ☐ Hearing Impairment
- ☐ Visual Impairment
- ☐ Mobility/Orthopedic Impairment
- ☐ Other
- ☐ None

Citizenship: (Choose one) ☐ U.S. Citizen ☐ Permanent Resident ☐ Other non-U.S. Citizen

Check here if you do not wish to provide any or all of the above information (excluding PI/PD name): ☒

REQUIRED: Check here if you are currently serving (or have previously served) as a PI, co-PI or PD on any federally funded project ☐

Ethnicity Definition:
Hispanic or Latino. A person of Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race.

Race Definitions:
American Indian or Alaska Native. A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment.
Asian. A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
Black or African American. A person having origins in any of the black racial groups of Africa.
Native Hawaiian or Other Pacific Islander. A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
White. A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

WHY THIS INFORMATION IS BEING REQUESTED:

The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PIs/PDs. To gather information needed for this important task, the proposer should submit a single copy of this form for each identified PI/PD with each proposal. Submission of the requested information is voluntary and will not affect the organization's eligibility for an award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information recieved from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. (The exceptions are the PI/PD name and the information about prior Federal support, the last question above.)

Collection of this information is authorized by the NSF Act of 1950, as amended, 42 U.S.C. 1861, et seq. Demographic data allows NSF to gauge whether our programs and other opportunities in science and technology are fairly reaching and benefiting everyone regardless of demographic category; to ensure that those in under-represented groups have the same knowledge of and access to programs and other research and educational oppurtunities; and to assess involvement of international investigators in work supported by NSF. The information may be disclosed to government contractors, experts, volunteers and researchers to complete assigned work; and to other government agencies in order to coordinate and assess programs. The information may be added to the Reviewer file and used to select potential candidates to serve as peer reviewers or advisory committee members. See Systems of Records, NSF-50, "Principal Investigator/Proposal File and Associated Records", 63 Federal Register 267 (January 5, 1998), and NSF-51, "Reviewer/Proposal File and Associated Records", 63 Federal Register 268 (January 5, 1998).

List of Suggested Reviewers or Reviewers Not To Include (optional)

SUGGESTED REVIEWERS:
Not Listed

REVIEWERS NOT TO INCLUDE:
Not Listed

COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 09-29

PD 06-7231 **08/13/09**

FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.)

OCI - STRATEGIC TECHNOLOGIES FOR CI

FOR NSF USE ONLY

NSF PROPOSAL NUMBER

0943725

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
06/01/2009	5	01110000 OCI	7684	072933393	06/03/2009 10:01am S

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)

951642394

SHOW PREVIOUS AWARD NO. IF THIS IS
☐ A RENEWAL
☐ AN ACCOMPLISHMENT-BASED RENEWAL

IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? YES ☐ NO ☒ IF YES, LIST ACRONYM(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE

University of Southern California

AWARDEE ORGANIZATION CODE (IF KNOWN)

0013284000

ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE

University Park
Los Angeles, CA 90089-1147

NAME OF PERFORMING ORGANIZATION, IF DIFFERENT FROM ABOVE

PERFORMING ORGANIZATION CODE (IF KNOWN)

ADDRESS OF PERFORMING ORGANIZATION, IF DIFFERENT, INCLUDING 9 DIGIT ZIP CODE

IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions)
☐ SMALL BUSINESS ☐ MINORITY BUSINESS
☐ FOR-PROFIT ORGANIZATION ☐ WOMAN-OWNED BUSINESS

☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE

TITLE OF PROPOSED PROJECT **STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ 1,611,038	36 months	09/01/09	

CHECK APPROPRIATE BOX(ES) IF THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW

☐ BEGINNING INVESTIGATOR (GPG I.G.2)
☐ DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C.1.e)
☐ PROPRIETARY & PRIVILEGED INFORMATION (GPG I.D, II.C.1.d)
☐ HISTORIC PLACES (GPG II.C.2.j)
☐ EAGER* (GPG II.D.2) ☐ RAPID** (GPG II.D.1)
☐ VERTEBRATE ANIMALS (GPG II.D.6) IACUC App. Date _____
PHS Animal Welfare Assurance Number _____

☐ HUMAN SUBJECTS (GPG II.D.7) Human Subjects Assurance Number _____
Exemption Subsection _____ or IRB App. Date _____
☐ INTERNATIONAL COOPERATIVE ACTIVITIES: COUNTRY/COUNTRIES INVOLVED (GPG II.C.2.j)
_____
☐ HIGH RESOLUTION GRAPHICS/OTHER GRAPHICS WHERE EXACT COLOR REPRESENTATION IS REQUIRED FOR PROPER INTERPRETATION (GPG I.G.1)

PI/PD DEPARTMENT
Information Sciences Institute

PI/PD FAX NUMBER
310-823-6714

PI/PD POSTAL ADDRESS
4676 Admirality Way, Suite 1001

Marina Del Rey, CA 90292
United States

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Electronic Mail Address
PI/PD NAME **Ewa Deelman**	**PhD**	**1997**	**310-448-8408**	**deelman@isi.edu**
CO-PI/PD **Burt Holzman**	**DPhil**	**2000**	**630-840-3922**	**burt@fnal.gov**
CO-PI/PD **Frank Wuerthwein**	**PhD**	**1995**	**858-534-3648**	**fkw@ucsd.edu**
CO-PI/PD				
CO-PI/PD				

Electronic Signature

CERTIFICATION PAGE

Certification for Authorized Organizational Representative or Individual Applicant:

By signing and submitting this proposal, the Authorized Organizational Representative or Individual Applicant is: (1) certifying that statements made herein are true and complete to the best of his/her knowledge; and (2) agreeing to accept the obligation to comply with NSF award terms and conditions if an award is made as a result of this application. Further, the applicant is hereby providing certifications regarding debarment and suspension, drug-free workplace, and lobbying activities (see below), nondiscrimination, and flood hazard insurance (when applicable) as set forth in the NSF Proposal & Award Policies & Procedures Guide, Part I: the Grant Proposal Guide (GPG) (NSF 09-29). Willful provision of false information in this application and its supporting documents or in reports required under an ensuing award is a criminal offense (U. S. Code, Title 18, Section 1001).

Conflict of Interest Certification

In addition, if the applicant institution employs more than fifty persons, by electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative of the applicant institution is certifying that the institution has implemented a written and enforced conflict of interest policy that is consistent with the provisions of the NSF Proposal & Award Policies & Procedures Guide, Part II, Award & Administration Guide (AAG) Chapter IV.A; that to the best of his/her knowledge, all financial disclosures required by that conflict of interest policy have been made; and that all identified conflicts of interest will have been satisfactorily managed, reduced or eliminated prior to the institution's expenditure of any funds under the award, in accordance with the institution's conflict of interest policy. Conflicts which cannot be satisfactorily managed, reduced or eliminated must be dislosed to NSF.

Drug Free Work Place Certification

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Drug Free Work Place Certification contained in Exhibit II-3 of the Grant Proposal Guide.

Debarment and Suspension Certification

(If answer "yes", please provide explanation.)

Is the organization or its principals presently debarred, suspended, proposed for debarment, declared ineligible, or voluntarily excluded from covered transactions by any Federal department or agency? Yes ☐ No ☒

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant is providing the Debarment and Suspension Certification contained in Exhibit II-4 of the Grant Proposal Guide.

Certification Regarding Lobbying

The following certification is required for an award of a Federal contract, grant, or cooperative agreement exceeding $100,000 and for an award of a Federal loan or a commitment providing for the United States to insure or guarantee a loan exceeding $150,000.

Certification for Contracts, Grants, Loans and Cooperative Agreements

The undersigned certifies, to the best of his or her knowledge and belief, that:

(1) No federal appropriated funds have been paid or will be paid, by or on behalf of the undersigned, to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with the awarding of any federal contract, the making of any Federal grant, the making of any Federal loan, the entering into of any cooperative agreement, and the extension, continuation, renewal, amendment, or modification of any Federal contract, grant, loan, or cooperative agreement.

(2) If any funds other than Federal appropriated funds have been paid or will be paid to any person for influencing or attempting to influence an officer or employee of any agency, a Member of Congress, an officer or employee of Congress, or an employee of a Member of Congress in connection with this Federal contract, grant, loan, or cooperative agreement, the undersigned shall complete and submit Standard Form-LLL, ''Disclosure of Lobbying Activities,'' in accordance with its instructions.

(3) The undersigned shall require that the language of this certification be included in the award documents for all subawards at all tiers including subcontracts, subgrants, and contracts under grants, loans, and cooperative agreements and that all subrecipients shall certify and disclose accordingly.

This certification is a material representation of fact upon which reliance was placed when this transaction was made or entered into. Submission of this certification is a prerequisite for making or entering into this transaction imposed by section 1352, Title 31, U.S. Code. Any person who fails to file the required certification shall be subject to a civil penalty of not less than $10,000 and not more than $100,000 for each such failure.

Certification Regarding Nondiscrimination

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative is providing the Certification Regarding Nondiscrimination contained in Exhibit II-6 of the Grant Proposal Guide.

Certification Regarding Flood Hazard Insurance

Two sections of the National Flood Insurance Act of 1968 (42 USC §4012a and §4106) bar Federal agencies from giving financial assistance for acquisition or construction purposes in any area identified by the Federal Emergency Management Agency (FEMA) as having special flood hazards unless the:

(1) community in which that area is located participates in the national flood insurance program; and
(2) building (and any related equipment) is covered by adequate flood insurance.

By electronically signing the NSF Proposal Cover Sheet, the Authorized Organizational Representative or Individual Applicant located in FEMA-designated special flood hazard areas is certifying that adequate flood insurance has been or will be obtained in the following situations:

(1) for NSF grants for the construction of a building or facility, regardless of the dollar amount of the grant; and
(2) for other NSF Grants when more than $25,000 has been budgeted in the proposal for repair, alteration or improvement (construction) of a building or facility.

AUTHORIZED ORGANIZATIONAL REPRESENTATIVE	SIGNATURE	DATE
NAME **Elizabeth L Hisserich**	**Electronic Signature**	**Jun 1 2009 5:38PM**

TELEPHONE NUMBER	ELECTRONIC MAIL ADDRESS	FAX NUMBER
213-740-6058	**ehisseri@usc.edu**	

* EAGER - EArly-concept Grants for Exploratory Research
** RAPID - Grants for Rapid Response Research

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Although much work has been done in developing the national cyberinfrastructure in support of science, there is still a gap between the needs of the scientific applications and the capabilities provided by the resources. Leadership-class systems are optimized for highly-parallel, tightly coupled applications. Many scientific applications, however, are composed of a large number of loosely-coupled individual components, many with data and control dependencies. Running these complex, many-step workflows robustly and easily still poses difficulties on today's cyberinfrastructure. One effective solution that allows applications to efficiently use the current cyberinfrastructure is resource provisioning. Today, the Compact Muon Solenoid (CMS) and other high-energy physics applications are using the glideinWMS system to dynamically provision resources as CMS jobs are being sent to the Open Science Grid. On the other hand Corral is being used to provision TeraGrid resources ahead of the execution of large-scale earthquake science workflows such as CyberShake. While the two provisioning systems have been developed independently they use a common underlying technology, Condor Glidein, and they provide applications with complementary capabilities.

Objective: The goal of this work is to integrate the two resource provisioning systems, to provide a common job interface to the two national cyberinfrastructures in the US (OSG and the TeraGrid), and to increase the number of applications using the systems to make efficient use of campus and national computing. The system will provide both a priori and dynamic provisioning capabilities, where resources can be reserved explicitly before application execution as well as implicitly as the application jobs enter the job management system. The effects of such provisioning strategies will be tracked via improved monitoring solutions integrated with the system. In addition to OSG and TeraGrid, this work will extend the provisioning capabilities to virtual environments such as those delivered by commercial and science clouds. New service administration and monitoring capabilities will be provided as part of this work. Summary statistics will provide information about system usage at the application and resource-levels.

Intellectual Merit The proposed system will provide a robust and scalable resource provisioning capability that will bridge heterogeneous, distributed cyberinfrastructures, making it easier for scientists with diverse computational requirements to efficiently leverage the available computing power and improve their overall productivity. The system is lightweight in that it resides in the application domain and has limited software requirements on the cyberinfrastructure such as remote access and outgoing network connectivity for cluster worker nodes. The proposed overlay-based architecture insulates researchers from infrastructure and resource failures and provides central control on execution policies over distributed resources. Efficient use of computational resources is also achieved by the integrated system by occupying resources as they become available and releasing those resources when they are no longer needed.

Broader Impact The integrated and extended system will support a broad spectrum of applications in the scientific domain ranging from workflows composed of interdependent tasks to applications composed of a large number of independent jobs to loosely coupled parallel applications. In addition to the current users of glideinWMS and Corral, the proposed system can be leveraged by applications in astronomy, gravitational-wave physics, fusion, and many others. Outreach will be done throughout the life of the project via established engagement, education, and dissemination avenues provided by OSG and TeraGrid. The software will be open source, available online and distributed through the well-established and broadly deployed Virtual Data Toolkit (VDT), thus increasing the reach of the proposed system to new communities. The software will be built and tested on the National Middleware Initiative (NMI) testbed on a large number of operating systems making it available to a broad range of domain researchers. The human resources necessary for the development of the proposed system will be selected in the respect of the institutional commitment to the integration of diversity in the workforce. Education of undergraduate and graduate students through involvement in related research projects will broaden the understanding of latest distributed system technologies and potentially spark new research endeavors.

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.B.2.

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

C 1. Objective: Integrated, robust, scalable, flexible resource provisioning for science applications

Vision The resource provisioning system proposed here will enable researchers to increase their scientific productivity and discovery potential through integrated, transparent management of small to very large ensembles of a diverse mix of computational jobs executed across the heterogeneous resources accessible through the US National Cyberinfrastructure.

Motivation The blueprint for a National Cyberinfrastructure [1, 2] put forward the vision of interdisciplinary collaborations of domain science researchers solving vital problems of our times via complex sets of simulation, data acquisition, and data analysis across heterogeneous computing resources ranging from locally owned clusters to leadership class machines. Today, these resources are presented to scientists as being distinct from each other, each with its own particular policies and access mechanisms. In this work, we propose to provide a software layer that creates an integrated cyberinfrastructure composed of resources governed by use policies ranging from local ownership to national allocations, to opportunistic use, to on-demand provisioning. This software layer thus provisions resources from a number of existing execution environments and makes them available, through a uniform interface, to scientific applications. It is generic in the sense that it is domain science independent and can be used by a broad set of applications, ranging from large numbers of single processor jobs to complex workflows.

In addition to presenting a uniform interface, resource provisioning can also improve application performance by reducing the overheads present in scheduling jobs on the current cyberinfrastructure. Most high-performance computing systems provide an execution model based on batch scheduling, where jobs are held in a queue until they can be matched with resources for execution. These systems provide only best-effort service and are typically configured to maximize resource utilization and not throughput. As a result, these systems often impose significant scheduling and queuing overheads on application jobs. For applications that consist of a single parallel job, or a few independent jobs, this model works well because the overheads do not contribute significantly to the total runtime of the application. For high-throughput workflows with large numbers of tasks, these overheads have a detrimental impact on performance, because delays increase at best linearly with the number of workflow tasks.

Goals **The goal of this work is to develop and deploy integrated, robust, and scalable resource provisioning services that supports a broad set of domain application workflow and workload execution environments. The aim is to integrate and enable these services across local and distributed computing resources, the major national cyberinfrastructure providers, Open Science Grid [3] and TeraGrid [4], as well as emerging commercial and community cloud environments. Additional user communities will be engaged and supported in their deployment and adoption of the technologies.**

The proposed system extends and reuses existing solutions that are already proven in production. Both Corral [5] and glideinWMS [6] currently operate as standalone resource provisioning systems. GlideinWMS was initially developed to meet the needs of the CMS (Compact Muon Solenoid) experiment [7] at the Large Hadron Collider (LHC) at CERN. It generalizes a Condor GlideIn [8] system developed for CDF (The Collider Detector at Fermilab) [9] and first deployed for production in 2003. It has been in production across the Worldwide LHC Computing Grid (WLCG), with major contributions from the Open Science Grid (OSG) in support of CMS for the past six months, and is being adopted for user analysis in time for data acquisition in October 2009. GlideinWMS also has been adopted by the CDF and MINOS [10] experiments, and is being evaluated by the DZero [11] experiment. GlideinWMS has been used in production with 8,000 concurrently running jobs, totaling more than six million jobs executed over the last year, consuming more than 2000 CPU years.

Corral [5], a tool developed to complement the Pegasus Workflow Management System [12, 13] was recently built to meet the needs of workflow-based applications running on the TeraGrid. It is being used today by the Southern California Earthquake Center (SCEC) CyberShake application [14, 15]. In a period of 10 days in May 2009, SCEC used Corral to provision a total of 33,600 cores and used them to execute 50 workflows, each containing approximately 800,000 application tasks, which corresponded to 852,120 individual jobs executed on the TeraGrid Ranger system. The 50-fold reduction from the number of workflow tasks to the number of jobs is due to job-clustering features within Pegasus designed to improve overall performance for workflows with short duration tasks [16].

The proposed extended integrated system, **CorralWMS**, will retain the identity of both systems to maintain backward compatibility for existing applications that use either Corral or glideinWMS, while new applications will benefit from the expanded set of provisioning capabilities of both systems combined. As with the current solutions, job management on the provisioned resources will be done via Condor [17] which allows for the customization of job scheduling policies. CorralWMS will support the management of up to 40,000 single-processor, possibly interdependent, jobs in dynamically provisioned scenarios, and enable the allocation and use of multiple resources for workflow-based and large-scale multi-processor computations. All stakeholders of the current systems (CMS, CDF, DZero, Minos, SCEC, etc.) will benefit from the increased capabilities of resource provisioning. We attach letters of collaboration from the CMS and SCEC projects describing their current reliance on glideinWMS and Corral respectively and the benefit of the proposed solution to their applications. Other communities already using OSG and TeraGrid, in particular those being supported by the OSG Engagement effort as well as others already looking for improved job management technologies, will be helped through adoption and support of the new system. We attach a letter of collaboration from the OSG Engagement group expressing interest in the new integrated system. In particular, communities that run parallel applications, such as plasma physics [18] or bio-informatics, will be able to concurrently schedule inter-dependent jobs on opportunistic resources, using the proposed features of pre-allocation.

This new system combines the benefits of the proactive provisioning performed by Corral with the dynamic resource provisioning performed by glideinWMS. In addition, the proposed program of work will extend the existing provisioning options to include cloud computing resources, which are of interest to the current stakeholder communities. Finally, the monitoring capabilities will be improved—extending the ability to monitor overall system behavior (the provisioning being performed) as well as the jobs making use of these resources.

In this work, we will leverage and contribute to the current activities of the OSG community engagement groups by supporting both OSG and TeraGrid resources and users directly and through the emerging OSG-TG job gateways. We will actively continue to support current users in earthquake science and high energy physics, as well as finding and engaging new users from the research community. Our new, extended, solution will thus be presented and tailored to the needs of diverse communities.

Developments Integration of the two existing systems will offer users the means to use both provisioned resources and "just-in-time" available resources for a single workload suite. It will provide a common interface to support provisioning of new types of resources, in particular clouds. The proposed development will benefit from the strengths of each existing system and will provide missing components for service administration and monitoring. Common monitoring capabilities, tailored views for real-time monitoring, and comprehensive generation of summary statistics will be developed.

Many science communities, as well as the cyberinfrastructure service providers, are cognizant of the potential opportunities offered by cloud computing resources – both commercial and community that are based on virtualization technologies. We will extend the integrated provisioning and job management system to provision and configure cloud resources for use by the general set of science applications.

Collaborations and Dependencies: Corral and glideinWMS make significant use of existing Condor components. The Condor Project is committed to working with us to support the integrated system and collaborating to ensure the most effective technical solutions are implemented (please see attached letter

of collaboration from the Condor project). The Open Science Grid will contribute software packaging and distribution support through the Virtual Data Toolkit (VDT), provide resources on the overlay test systems (please see attached letter of collaboration from OSG), and contribute scalability and performance testing as part of the extensions program of work at UCSD.

Software build, test, releases, and licensing: The software proposed here will use the National Middleware Initiative (NMI) build and test system. The software will be packaged and distributed through the OSG VDT [19]. VDT is funded through the Open Science Grid, thus by the National Science Foundation and the Department of Energy, and as such has a broad impact on a great number of scientific applications. The program of work will include development of an expanded test harness that will test the individual system components as well as the end-to-end system. Our existing software is open source. The integrated system will also be open sources and use a software license from www.opensource.org.

C 2. Intellectual Merit and Broader Impact

Intellectual Merit: The proposed work will have an impact in five main areas:

1. **Providing an interface to a number of heterogeneous cyberinfrastructure resources**, which would make it easier for scientists to leverage the existing campus and national cyberinfrastructure. Thus it will advance the vision of computational thinking that is becoming the foundation of 21st century research [20]. While there is a wealth of resources available across local campuses, national service providers like OSG and TeraGrid, federated international providers such as WLCG, and across cloud computing resources, it can be difficult for user communities to manage these heterogeneously administered resources effectively according to their diverse internal policies and needs. The overlay-based architecture of CorralWMS, which builds a virtual resource pool out of existing resources, insulates the scientists from resource and infrastructure failures (overlays are affected rather than application tasks), and provides a central control for scientific workloads (queue priorities, access policies, low-latency job control) instead of distributed control at many heterogeneous resources.
2. **Providing a resource provisioning capability that can support a number of different types of scientific applications** including single processor codes, parallel applications, and scientific workflows. The capability would also potentially result in better overall application performance and thus increase the scientists' productivity. Provisioning resources from disparate infrastructures and creating a highly dynamic community-dedicated resource overlay to be used by applications and services (such as job schedulers) are key. Software-based interpretation of usage policies can enable optimal use and sharing of the available resources. Our program of work makes contributions in all these key areas. And, since our program of work brings together technologies already in production use by multiple communities, we have confidence that the added capabilities will provide practical benefit and added value quickly.
3. **Providing an efficient and robust job management capability** that can automatically select and make easy use of the appropriate provisioned resources, thus providing automation and improving fault tolerance for the applications. Providing resource provisioning for scientific applications has also shown significant improvements in application performance [21, 22].
4. **Leveraging the emerging cloud resources** to support dynamic, on-demand classes of applications such as those that perform weather prediction. We envision that the seamless integration of resources from existing Grid infrastructures and emerging cloud environments will provide a high degree of flexibility in the allocation of computing cycles. Today, such integration is in its infancy, as are some of the cloud technologies, and it is far from seamless. The integration problem is made more complex by the model of Grids and clouds as complementary technologies: Grids allow uniform access to distributed resources, while clouds provide uniform (customizable) execution environments. Because of this complementary nature, the integration of these technologies is key to achieve the true goal of distributed systems: uniform access to uniform environments. Considering that the focus of our work is resource provisioning for scientific communities, we believe that our program is well placed to give a strong contribution to the Grid /cloud integration problem.

Broader Impact: The two infrastructures used, OSG and TeraGrid, are major distributed resource providers contributing to the US national cyberinfrastructure today. Users of these facilities span a large number of scientific applications ranging from astronomy [23], biology [24, 25], ecology [26] earthquake science [27, 28], gravitational-wave physics [29], high-energy physics [30-33], and many others. These diverse communities use the ensemble of resources with different methods and patterns based on their domain and application needs, including single-processor jobs, multi-processor MPI-based codes, and workflows that place additional data and control dependencies between application jobs. All these applications can potentially benefit from the proposed provisioning system by leveraging a number of different resources to do science. Our community engagement efforts are geared towards reaching out to new communities, making them aware of the new capabilities, and assuring that our work directly benefits their computation needs.

Broader Impact: Integration of Research and Education: We will actively promote the use of this system for teaching and learning, showing how it makes it easier for students to make use of a disparate cyberinfrastructure and to leverage it for their domain research. Computer science students can develop and test algorithms that can be used to direct the behavior of the system: how many resources to acquire, when, etc. We will also make de-identified traces of system usage available to the community (the number and types of provisioning requests, the number and basic characteristics of application jobs). Such real-life resource utilization patterns are widely considered much more valuable than results from Grid utilization simulations. In turn, these data can help further tune these simulators to achieve a greater level of realism. Furthermore, data on resource utilization patterns can be used by students to develop new resource provisioning and task scheduling algorithms. The provisioning infrastructure will also enable Ph.D. students to study heterogeneous cyberinfrastructures.

We will continue to work with students to enable them to use state-of-the-art tools for their research. Dr. Deelman will also integrate the technologies under this proposal as part of the curricula at USC. Additionally, graduate students working on research projects either through directed research coursework or graduate student assistantships will benefit from the proposed infrastructure.

OSG plans to include glideinWMS as one module in the e-learning environment being developed as an extension to the Grid Schools which have been conducted over the past five years [34]. The next step will be to include CorallWMS in that educational setting.

Broader Impact: Result dissemination: The main vehicles for dissemination of results of this proposal will be a series of software releases and user adoption of the technologies. The releases will be supported by the NMI test and build infrastructure. The software and documentation will be made available on the project website, via a Subversion code repository, and as mentioned above as part of VDT. Users of the technologies will provide natural dissemination of their utility and value. Additionally, we will disseminate the results of this work through tutorials and presentations at conferences, and in conference and journal publications.

Broader Impact: Integrating Diversity: Dr. Ewa Deelman is part of the USC Women in Science and Engineering (WiSE) program. WiSE aims at increasing the participation of women in science and engineering at various stages of their careers, from middle school students through Ph.D. programs to faculty. As part of WiSE, Dr. Deelman mentors women in the Computer Science graduate program and participates in meetings of the female faculty in the USC School of Engineering. Dr. Deelman also mentors undergraduate women in Computer Science through the CRA-W Summer mentorship program. As part of these internships, students can learn about the challenges of scientific applications and new technologies such as those proposed here. These internships can also resource in research publications [35]. Fermilab fosters diversity in its working environment through the mission of the Diversity Council. The council fosters organizational equity through programs carefully designed to increase the diversity of the Laboratory and to increase the participation of employees. The Council develops, implements, and maintains strategic programs with established goals for the Laboratory. It will organize teams that will develop initiatives in three areas: workforce recruitment, retention, and development; educational

outreach; and community involvement and outreach. The project will be able to employ these initiatives to perform educational outreach.

Appropriateness of the STCI program: The proposed work fits under STCI in several ways:

- There are no other calls for proposals in the Office of Cyberinfrastructure (or NSF in general) that would fit the scope of the proposed work (for example the SDCI program is no longer open to new proposals).
- The proposed work can potentially transform a number of areas of research: astronomy, biology gravitational-wave physics, earthquake science, high-energy physics, and others.
- The proposed work will not generate outcomes that are under development under the PI's current funding. Existing projects on the development of Pegasus focus on the workflow system, not the provisioning capabilities. Current work within the glideinWMS project is focused on solely provisioning opportunistic resources such as OSG. The OSG project scope does not include software development.
- This proposal describes integration of two complementary technologies to make a more general solution, as well as the development, testing, and maintenance of new capabilities for the future.
- The proposed work is applicable to a number of different science communities and can be integrated with existing workflow and application management systems.

C 3. Proposers' Qualifications

Dr. Ewa Deelman (PI) is a leader in the field of scientific workflows. Dr. Deelman's software, Pegasus, is being used in production today by a number of projects including Montage (astronomy) [23, 36], SCEC (earthquake science) [27, 28, 37], LIGO (gravitational-wave physics [29], Epigenetic Center at USC, and others. Besides heading the Pegasus project since its beginning in 2001, she has extensive collaborations with both computer and domain scientists. Recently, she co-edited the book on "Workflows for e-Science"[38], the first comprehensive treatment of scientific workflows, which includes contributions from many application scientists as well as most of the scientific workflow researchers in the US and Europe. In 2006 Dr. Deelman also co-chaired an NSF-funded Workshop on Challenges of Scientific Workflows which brought together thirty researchers in the field of scientific workflows to discuss current challenges in the area of applications, workflow representation, dynamic workflows, and workflow systems.

Dr. Burt Holzman (co-PI) has a successful track record of producing production software for scientific research. From 2002 through 2005, Dr. Holzman was the head of computing for the PHOBOS experiment at the Relativistic Heavy Ion Collider (RHIC), leading software development and planning and coordinating large-scale operations at the RHIC computing facility. As the leader of support efforts for CMS on the Open Science Grid, Dr. Holzman has a wide range of expertise with grid technologies, and has served as release co-coordinator for Open Science Grid releases in the past. Dr. Holzman currently leads several projects contributing to the Virtual Data Toolkit and Open Science Grid software stacks, including the Generic Information Provider and the OSG/EGEE Integration Interoperability activity. Dr. Holzman is the head of the glideinWMS project, which is currently in use by a number of large scientific collaborations, including CMS, CDF, and MINOS.

Dr. Frank Würthwein (co-PI) and Igor Sfiligoi have worked on development, deployment, and operations of glidein-based systems. They started with a system for CDF that reached a scale of several thousand simultaneously running jobs, serving hundreds of physicists for many years. In 2006, Sfiligoi extrapolated from the lessons learned with CDF, and created glideinWMS as a generic resource provisioning tool. Würthwein led the commissioning of the global CMS compute and data grid, and is presently creating the data analysis operations task within international CMS. He is thus part of the international CMS management team. He is presently co-leading the Applications area in the Open Science Grid, is a member of the OSG Executive Team, and was the founding Chair of the OSG

Executive Board. Sfiligoi is presently leading the "Scalability, Reliability, and Usability" area within OSG, in addition to contributing to glideinWMS development and operational security in OSG.

Dr. Gabriele Garzoglio (Sr. Personnel) works as a computer professional at Fermilab. He leads the Open Science Grid group, a team of Software Engineers that provides middleware solutions for large distributed systems. In recent years, Dr. Garzoglio has focused on data and workload management for high energy physics applications, as well as resource selection and authorization.

Community Engagement:

All participants have a history of community engagement and technical collaboration. They have many years of experience working collaboratively in support of data intensive scientific communities. Most recently they have worked together to support stakeholders using the Open Science Grid infrastructure.

Dr. Deelman has worked with a number of science projects, and she will continue to work with these communities to make the work relevant. Dr. Holzman has worked extensively with researchers on CMS in order to enable their work to continue on the OSG and WLCG. He has been integral to the data analysis and user support of the Nuclear and Particle Physics Experiments of which he has been a member. He has made significant contributions to other communities through contributions to the OSG program of work in the areas of integrated system testing, information services and interoperability between grid infrastructures. Würthwein has a history of engagement with a number of communities within the OSG. His group at UCSD is leading the virtual organization coordination effort within OSG, and thus working with all scientific communities within OSG other than the three primary stakeholders ATLAS, CMS, and LIGO. In addition, he is responsible for coordinating data analysis operations on the global CMS data grid. Deelman and Würthwein have worked together on engaging the LIGO community in the OSG and enabling them to run workflow-based applications on these cyberinfrastructure resources. Holzman, Sfiligoi, and Würthwein are working together within the context of both CMS and OSG on a variety of topics. All will leverage their current collaborations to disseminate the results of this work and gain community feedback.

Results from prior NSF Support: Dr. Deelman is the PI for OCI- 0722019 (SDCI-Pegasus), 9/1/2007-8/31/2010, $1,700,000. As part of this effort Dr. Deelman is continuing to improve, support, and maintain the Pegasus Workflow Management System that is used by a number of scientific communities. New workflow restructuring techniques as well as new workflow design capabilities were developed in response to community needs. Dr. Deelman was a co-PI of the NSF-funded Workshop on Scientific Workflow Challenges (WSW-06), (IIS- 0629361), 5/1/2006-10/31/2007, $44, 961, which examined workflow issues from the perspective of application and computer scientists and included 30 computer science and domain experts. The report and all workshop materials are available at: www.isi.edu/nsf-workflows06 and http://www.nsf.gov/od/oci/reports.jsp.

Dr. Würthwein is a Co-PI, and technical lead on the NSF funded DISUN grant (PHY-0533280). DISUN is a grant heavily focused on integration and operation of Cyberinfrastructure. Its $10 million over five years include 50% hardware and 50% personnel. Some of the 2008/09 DISUN highlights include commissioning of the CMS data grid [39], use of glideinWMS in CMS [40], glideinWMS scalability testing [41], Analysis Support Task Force in CMS [42], pseudo-interactive monitoring in distributed computing [43], the pilot way to grid resources [44], Commissioning of the CMS computing centers [45] and many others.

C 4. Resource Provisioning: Benefits, Existing Solutions and Current Applications

The proposed program of work aims to provide resource provisioning to allow the composition of community-dedicated "overlay" scheduling systems, dynamically composed of distributed resources, but uniformly manageable by the individual communities. The advantages of such an approach are multifold. First, once computing resources are provisioned, the delay for dispatching jobs to the infrastructure is short, resulting in fast turn-around, and, even more importantly, predictability of delays. Jobs do not need to wait in the local scheduling queues at the sites; rather they are assigned to the pre-allocated provisioned

resources. This type of scheduling is critical to the performance of workflow-based applications where queuing and scheduling delays accumulate as jobs in the workflow are released for execution to the cyberinfrastructure. Second, the scheduling policies for the jobs of the community rest with the community itself, instead of being a hard-to-model composition of the scheduling policies at the individual site schedulers. In particular, execution priorities can be controlled through the configuration of the overlay scheduling system. Third, systematic incompatibilities in the configuration of the underlying resources with community requirements do not result in failures of the user jobs. In fact, the provisioning system deployed for a given community, through the use of pilot jobs (see below), will not allocate resources that are incompatible with the operations of that community.

These benefits apply to both single and multi-processor applications. In addition, for the latter, resource provisioning allows for the usage of opportunistic resources for concurrent jobs. As discussed above, some communities (e.g. ITER) run loosely coupled concurrent jobs of small size (64 processors); therefore, they can efficiently use Ethernet-based computing clusters of the cyberinfrastructure, which are abundant in grids such as OSG. For these use cases, however, the OSG model of opportunistic usage of resource cannot be used without resource provisioning. In fact, only a small fraction of resources at each cluster are dedicated to opportunistic usage, and the probability of accessing all needed resources for concurrent jobs without reservation is very low, in most cases. This problem can be addressed by a provisioning system requesting concurrent blocks of resources at such clusters.

C 4.1 High-Energy Physics and the glideinWMS on OSG

The Compact Muon Solenoid (CMS) experiment is one of two general-purpose particle physics detectors constructed for the Large Hadron Collider at CERN. More than 2,000 physicists from more than 30 countries will use the data collected to search for new phenomena including the Higgs boson, extra dimensions, and supersymmetry. CMS uses a globally distributed computing model. The data is generated and first processed at CERN, and then transferred to the seven regional Tier-1 data centers around the world. The Tier-1 centers then skim the processed data for physics signals, perform further processing ("re-reconstruction") and serve them to Tier-2 centers for analysis.

CMS will regularly reprocess the data as it is collected during 2009 and 2010, culminating with a complete processing pass of $2.5e^9$ events stored in 3.8 PB of input files and 1.3 PB of output files. Tasks in the final pass are expected to consume eight hours of processing time individually and together will utilize approximately 800 integrated CPU-years. It is desired that the production managers be able to selectively prioritize re-reconstruction of different datasets based on physics priorities.

The Tier-1 centers will also execute high-priority, high-throughput tasks known as "skim jobs" against the 1.3 PB dataset. The skim workflow consumes a large amount of I/O with short jobs lasting 30-60 minutes. The use of glideinWMS greatly reduces the overall latency in running skim workflows. It is also important to ensure that skim workflows are properly prioritized with respect to re-reconstruction workflows. GlideinWMS provides centralized control across all seven Tier-1 centers, as shown in Figure 1. The Figure shows the number of glideins submitted over time to the Tier 1 sites (some of which are part of OSG) in response to user job submission.

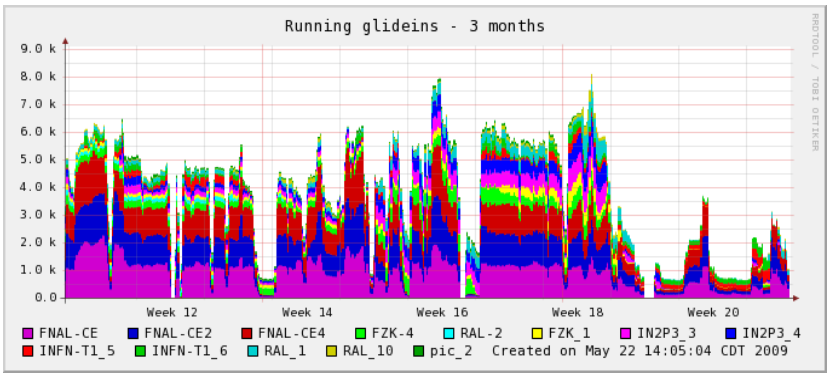


Figure 1: CMS T1 skim workflows over the past 3 months.

Roughly 8,000 cores are pledged for analysis at the 45-60 Tier 2s available worldwide. CMS has developed the CRAB tool to allow researchers to submit workflows to the grid. The length and scope of

analysis workflows are defined by the physics working groups and the scientists performing them. Their length is highly variable, lasting from minutes to days. GlideinWMS provides the ability to centrally manage analysis priorities based on scientific need, and once again significantly lowers latency in the case of very short analysis tasks. Also, with a large number of available resources for analysis, glideinWMS insulates the scientist from site issues, and submission efficiency to the grid remains very high.

GlideinWMS is based on the Condor glidein paradigm—dynamically creating virtual-private Condor pools out of Grid resources. Glidein [46] is a multi-level scheduling technique where pilot jobs called "glideins" are used to start Condor worker daemons on remote cluster nodes. Upon startup, the worker daemons join a Condor pool administered by the user (i.e. a personal cluster) where they are used to execute application jobs. The glideinWMS is composed of five elements, as shown in Figure 2:

1. A resource-less Condor pool, composed of a **Condor scheduler**, a **Condor collector** and a **Condor negotiator** (not shown in the figure): These services represent the primary interface to the end-user. The collector keeps track of the resources as they become available. The scheduler handles the user jobs. The negotiator handles the VO policies.
2. A **Glidein Factory**: This service adds resources to the Condor pool, by submitting glideins (described below) to Grid sites. The Glidein Factory contains the list of all the trusted Grid sites and their configuration. It periodically advertises this information to the WMS Collector (described below) and waits for requests from a VO Frontend (described below). It submits glideins via Condor-G to appropriate sites accordingly.
3. **A WMS Collector:** This is a message exchange service. It handles network traffic between VO Frontends and Glidein Factories, including handling authentication and authorization.

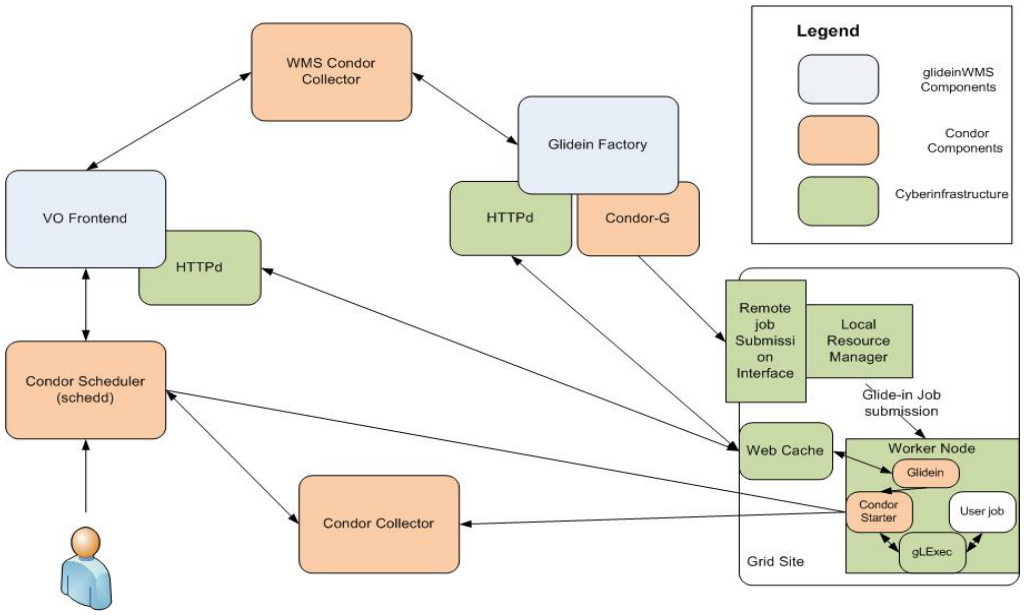


Figure 2: Architecture of glideinWMS showing Condor and Glidein components.

4. A **VO Frontend**: This service regulates the number of glideins sent to various Grid sites. The VO Frontend periodically probes the user collector for idle jobs and matches them against the information provided by a Glidein Factory. It uses this information to decide how many new glideins are needed and advertise this to the WMS Collector as requests for the Glidein Factory.
5. A set of **glideins**: These are the resource providers. They are scripts that configure and launch a Condor Starter. Once started on a Grid worker node, a glidein fetches the Condor binaries and a set of validation and configuration programs from the VO Frontend and the Glidein Factory. A site-local Web cache is used, if available, to minimize network traffic. The validation programs are executed

first, and if they all succeed, the configuration programs prepare the starter configuration file. Once the Condor Starter is launched, the glidein delegates all control to it. After the starter terminates, the glidein is responsible for the final cleanup. All network communication is strongly authenticated.

C 4.2 Earthquake Science and Corral on the TeraGrid

Researchers at the Southern California Earthquake Center (SCEC) use large-scale grid-based scientific workflows to perform seismic hazard research as a part of SCEC's program of earthquake system science research. The scientific goal of the SCEC CyberShake project is to calculate probabilistic seismic hazard curves for sites in Southern California. For each geographic site of interest, CyberShake includes two large-scale MPI calculations and approximately 840,000 embarrassingly parallel post-processing jobs. Thus, the CyberShake computational platform requires both high-performance computing (for the MPI jobs) and high-throughput computing (for the post-processing). To make a Southern California hazard map practical, time-to-solution per site needs to be short, on the order of 24-48 hours. This emphasis on reducing time-to-solution, rather than categorizing the system as a capability or capacity platform, pushes the CyberShake computational platform into the high-productivity computing category which is emerging as the key capability needed by science groups. In contrast to capability computing (a single, large job) and capacity computing (smaller, multiple jobs, often in preparation for a capability run), high productivity computing focuses on high throughput jobs with extremely short runtimes. The challenge is to minimize overhead and increase throughput to reduce end-to-end wallclock time.

In order to minimize the overheads, SCEC uses the Corral provisioning service to acquire the TeraGrid resources ahead of the workflow execution, and it uses the Pegasus WMS to restructure the workflow and manage the execution of the workflows on these resources. As a result of resource provisioning and workflow performance optimization, SCEC can run a CyberShake workflow consisting of over 800,000 tasks on 800 cores in less than 3 hours.

Corral, similar to gideinWMS, creates personal clusters by provisioning resources from grid sites using *glideins*. Corral was developed using a service-oriented architecture. Clients send provisioning requests to a grid service, which communicates with grid sites to allocate resources that fulfill the requests. The components of the system and the relationships between them are shown in Figure 3.

The Grid Site consists of a head node, several worker nodes, and a shared file system that can be accessed by all nodes. The head node hosts a Globus gatekeeper that accepts batch jobs, and a local resource manager (LRM) that matches jobs with resources.

The Corral Service accepts requests from clients, sets up the execution environment on the grid site, provisions resources using glideins, and cleans up files and directories created by the system. This service was developed using the Globus framework [47].

Condor is used to process service and application jobs, and to manage glidein workers. Corral submits

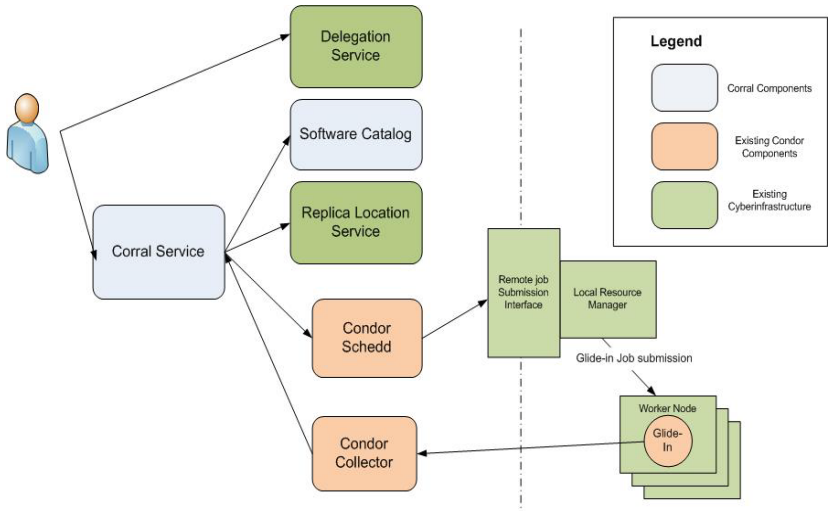


Figure 3: Corral Components.

glidein jobs to the grid site using the Condor-G facility in Condor, the glideins contact the Condor central manager to join the user's personal cluster, and application jobs are submitted to the Condor queue where they are matched to glideins for execution.

The Delegation Service is used to implement standard grid security mechanisms [48]. Clients send their security credentials to the Delegation Service where they are stored for later use by the Corral Service (for authentication when submitting jobs and transferring files to the grid site).

Staging Servers are file servers used to host bundles of executables and configuration files called *packages*. Each package contains a set of Condor worker node daemons for a different combination of Condor version, system architecture and operating system. Any file server that supports the HTTP(S), FTP, or GridFTP [49] protocols may be used as a staging server.

The Replica Location Service (RLS) [50] is used to map package names to staging servers.

The process used by the service to provision resources is divided into three phases: setup, provisioning, and cleanup. These phases correspond to jobs that are submitted by the service to the grid.

The Setup Job is submitted during the *setup phase* to prepare the site for glideins. The setup job runs an installer which determines the appropriate package to use for the site, looks up the package in RLS to determine which staging servers have it, and downloads the package from the first available staging server. It then creates an installation directory and a working directory on the shared file system, and unpacks the Condor binaries.

The Glidein Job is submitted during the *provisioning phase* to allocate worker nodes for the user's personal cluster. Glidein jobs generate a Condor configuration script and launch Condor worker daemons on each allocated node. The daemons are monitored by a special process and killed when the user's request is cancelled or expires.

The Cleanup Job is submitted during the *cleanup phase* to remove the working directories used by the glideins. It runs an uninstaller which removes all log files, configuration files, executables and directories created by the service.

Using this three-step process allows Condor executables to be staged once during the setup phase and reused for multiple requests during the provisioning phase. This precludes the transfer of binaries for each provisioning request and thereby reduces the provisioning overhead of the system.

C 4.3 Benefits of Provisioning resources for Workflow-based applications

We performed controlled experiments quantifying the benefits of performing workflow provisioning with Corral. Although we would have liked to perform the experiments on the TeraGrid using the SCEC application, running CyberShake without provisioning is not feasible because the TeraGrid resource providers often view the submission of 800,000 small jobs as a denial of service attack. Thus, we ran these tests on a high-performance cluster at ISI using the Montage astronomy application [36]. Montage workflows generate science-grade image mosaics of the night sky. The size of a workflow is based on the area of the sky (in square degrees) covered by the mosaic. The nodes in the cluster have 800MHz Pentium III processors and 1GB of memory. We used 2 different sized workflows for our evaluation: 1-degree (206 tasks) and 6-degree (6062 tasks). As can be seen in Figure 4, the runtime of the workflows using Corral was up to 78% less than the runtime using Globus (45% on average)—the default mode of execution without resource provisioning.

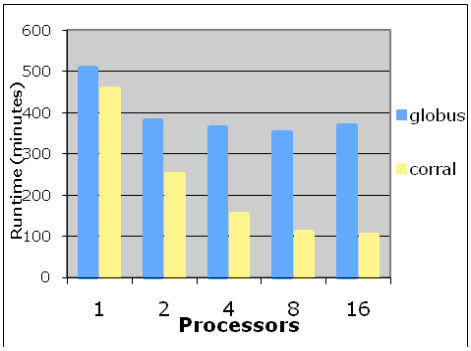


Figure 4: Comparison of the runtime of the workflows using Corral and Globus. Corral is up to 78% faster

When running on the existing cyberinfrastructure, such as the TeraGrid, site policies often dictate how many jobs a user can run at any one time. This poses obvious restrictions on large-scale workflow-based applications composed of single processor jobs (such as CyberShake, Montage, the

Epigenomic workflow, etc.). For such applications, provisioning a block of resources using a single job (as can be done with Corral) allows them to get their work done much faster. We illustrate this point with an Epigenomic workflow running on a USC cluster. The application consists of 2057 tasks that reformat, filter, map and merge DNA sequences. The majority (> 90%) of the runtime of this application is consumed by 512 tasks that each require approximately 2.5 hours to run. The experiments were executed on a 10GB cluster. The nodes we used had 2.3GHz AMD Opteron processors and 16GB memory. The cluster has two scheduling policies that affect the runtime of the genome application: *max_user_run*, and *resources_max.walltime*. The *max_user_run* policy prevents any single user from running more than 30 jobs concurrently. When using Globus to submit jobs, this policy prevents workflows with serial tasks from using more than 30 processors at a time. The *resources_max.walltime* policy prevents any single job from running for more than 24 hours. It is important to note that although workflows using Globus are limited to 30 processors by HPCC's *max_user_run* policy, workflows using Corral are not. Corral can use a single parallel glidein job to provision any number of processors. This allows Corral to allocate more processors to run the workflow than is possible using Globus. Figure 5 shows the runtime of the genomic workflow using 128, 256 and 512 processors. For this workflow 512 processors is the maximum parallelism achievable. Using Corral to provision 512 processors resulted in an order of magnitude lower runtime (211 minutes) compared with the best runtime using Globus (2349 minutes).

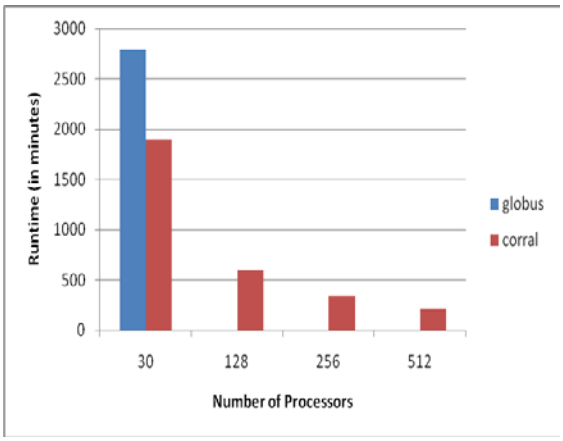


Figure 5: A Comparison of the Runtime of the Genomic Workflow Using 128, 256 and 512 processors via Globus and Corral.

C 5. CorralWMS: The Integrated Corral and glideinWMS System

The conceptual framework of the integrated system is shown in Figure 6. The *Domain Application* will be domain specific (e.g. CMS CRAB), either based on direct Condor submission, on general workflow tools such as Pegasus, or using portals such as the TeraGrid science gateways. The two *Resource Management components – provisioning and dynamic allocation* – provide alternate methods of provisioning. The provisioning resource management component is based on the static allocation of cores and/or nodes on a site, and the dynamic allocation component enables selection of resources based on determined availability – for example from a site database or information service. The integrated system enables both these methods to be *used across multiple architectural implementations of distributed resources*.

The software component architecture of the integrated system is shown in Figure 7. The user will be able to interact with the system in two ways. The user can submit a specific provisioning request, for example request 2,400 cores from Ranger (TG) or the user can rely on the system to provision resources as needed.

In the case of a specific request, the *Explicit Provisioning Service (EPS)* will place a provisioning request in a *Condor Collector (P),* which is used for communication between provisioning system components. Through the Condor advertising mechanism, the *Factory* is able to determine that provisioning needs to occur, and submits a *Condor glidein* to the desired cyberinfrastructure site. If such system is not specified, the *Factory* uses the *site information* to send the *Condor glideins* to the available resources. If the provisioning request cannot be satisfied in a given amount of time, the request may be broken up into smaller sized requests. For reasons of scalability and robustness, the *Factory* relies on a dedicated provisioning Condor scheduler (*schedd (P)*) to manage the provisioning requests. In addition, the system uses a separate application-dedicated Condor scheduler (*schedd (A)*) to manage the execution of application jobs on the provisioned resources. Once *Condor glideins* start running on the remote

resources, they communicate with *schedd (A),* and application jobs can now run directly on the remote system.

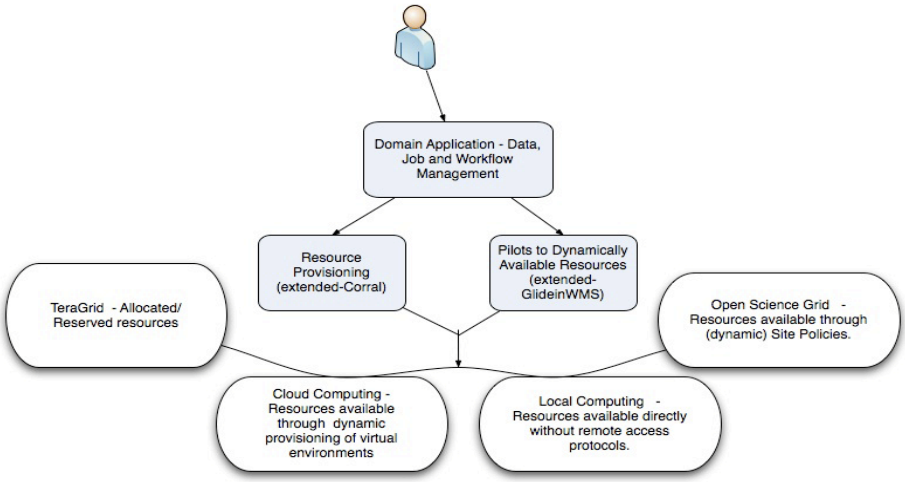


Figure 6: Integrated System Provides Access to Different Cyberinfrastructure Environments

If the user does not make explicit requests, but instead just submits jobs to schedd (A), the proposed system will function in a fully automated mode. The Automated Provisioning Service (APS) collects information about these jobs, and requests for resource provisioning are placed in the *Condor Collector (P)*. From there the provisioning occurs as before, with the Factory performing the provisioning.

Both EPS and APS also have the information about the applications needed by the user community. This information is delivered, through the Factory, to the glideins for proper application setup. The Factory also has the information about the Condor executables needed by the glideins; this information is provided to the glidein startup script. The delivery of user community applications and the Condor binaries could be either provisioned in advance, as currently done by Corral, or loaded on the fly, as currently done by glideinWMS, based on the provisioning system configuration or a user request.

The *Monitoring Service (P)* keeps information about all the elements of the provisioning system, and receives this information from the *Condor Collector (P)*. It records information about both requests for provisioning and the resources that were provisioned. The Monitoring information will be archived for a period of time. Cumulative statistics of service metrics will also be archived. We will provide user-friendly interfaces to the Monitoring Services so that users can easily visualize the behavior of both the provisioning system as well as the provisioned resources. This information is also important for system administrators so they can see whether resource utilization is appropriate. If the performance of the system needs to be tuned, either for the way that automatic provisioning is done or how sites are selected for resource provisioning, administrative interfaces will be provided.

The system will also provide a second *Monitoring Service (A)* that will gather and keep detailed information about the application jobs and the provisioned resources as seen by the users. The information will be archived as described above and will have a similar user interface. The reason for having a separate monitoring system is due to the fact that we want to allow a single provisioning system to serve multiple application job submission systems, and a single application job submission system to be served by multiple provisioning systems for maximum flexibility.

Although today we target mostly grid resources, the *Factory* in the new system will be extended to support cloud-based commercial environments such as those provided by Amazon EC2 [51] and science clouds such as those enabled by Eucalyptus [52] or Nimbus [53] software.

In this work we face several challenges. Some relate to assuring backward compatibility to the existing systems on which the applications rely. Some challenges arise from the inherent tension between implicit and explicit allocations. Jobs submitted to the system will either target pre-allocated resources or take

advantage of opportunistically provisioned resources—or both, in the case where the pre-allocated resources are not sufficient. We will rely on the Condor Matchmaking mechanisms [54] to assign jobs to the correct resources. Although the Condor Collector capabilities provide us with basic communications between the components, we will need to define new "protocols" to facilitate new modes of operation. Finally, cloud technologies are ever evolving as new capabilities are added. We will follow developments in that space and tailor our solutions to new environments and to our users' needs.

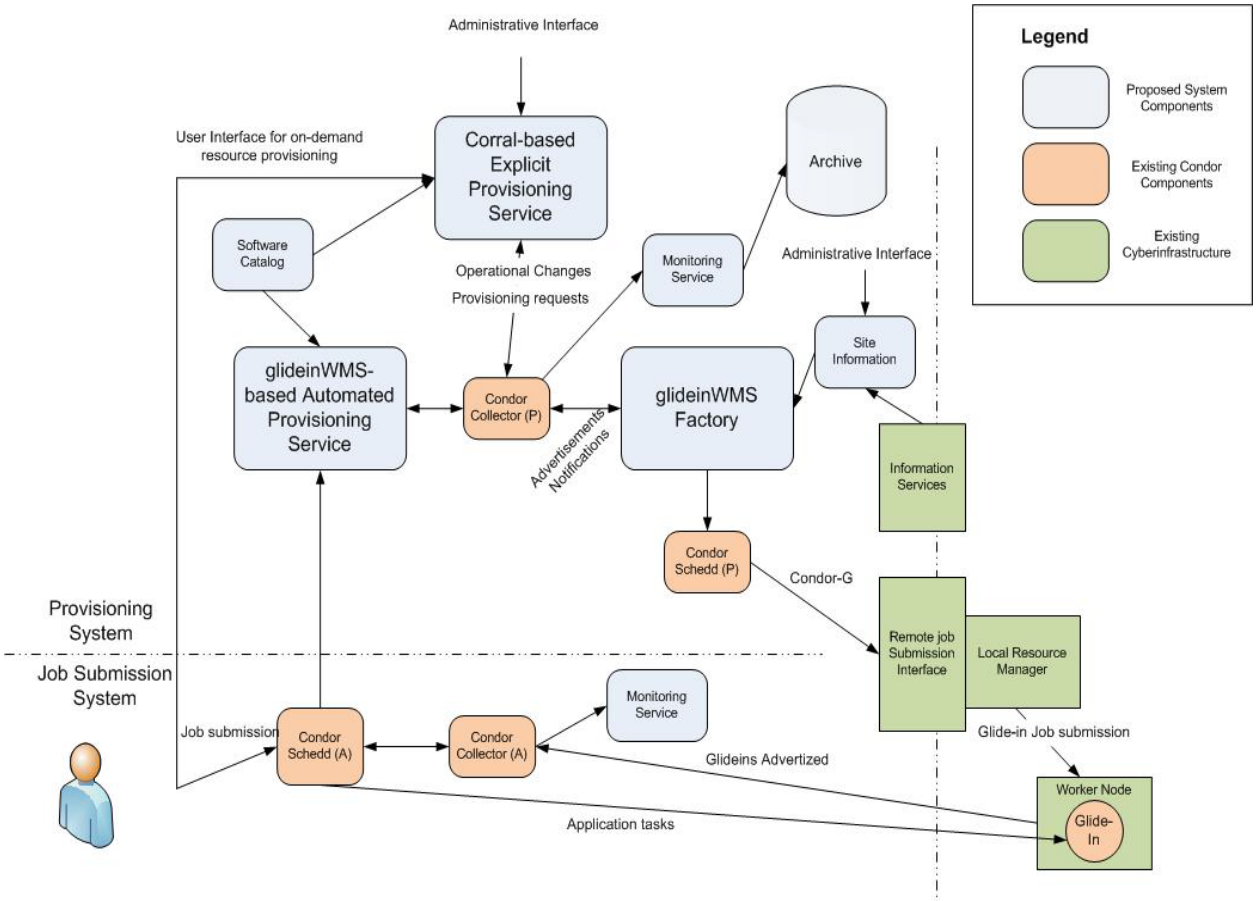


Figure 7: Architecture of the Integrated Corral-glideinWMS System.

C 6. Program of Work

The program of work allows for continued release and deployment of production quality software during the lifetime of the project. The work includes extending the use and support of the system to additional research communities. We provide a breakdown of the project goals in terms of annual deliverables and metrics. Institutions with primary responsibility are identified at the end of each task..

Year 1: Initial Integration of a Combined System and Development of Application-level Monitoring

Integration and Provisioning:

- Develop a common build and release structure that satisfies the current customer base of US CMS, HEP experiments, SCEC, and other Corral users. (USC, Fermi)
- Thoroughly test and validate the integrated system using existing test suites and applications. (ALL)

Development:

- Integrate the Corral frontend to the glideinWMS factory (USC)
- Develop dynamic switching between Corral and glideinWMS frontends (USC, Fermi)
- Support for multi-processor compute nodes in the glideinWMS environment. (Fermi)

- Port the Corral user request handler and factory interface to restful service. (USC)
- Design and develop a prototype of the application level monitoring system. (Fermi)
- Prototype collection of statistics on use of the software, collection and textual interpretation of error conditions to aid the domain application and end user. (USC, UCSD)

Collaboration and Community:

- Establish a collaborative workspace and develop the methods for collaborative development, software release and documentation. (Fermi, USC)
- Socialize the collaboration with the existing user base to enable continued trust and commitment to the software base. (ALL)

Outreach and Training: (ALL)

- Establish training methods and technologies.
- Contribute hands-on training for an example application at a SuperComputing tutorial, TeraGrid and OSG Grid School to understand the context and needs of these educator communities.

Metrics:

- Initial release of integrated system accepted by existing user communities.
- Number of jobs executing in parallel. Goal: >20,000.
- Number of jobs per day. Goal: >100,000
- Number of user communities. Goal: > 4 user communities
- Set a realistic metric for cloud computing use, based on understanding of existing user community needs, as well as cloud pricing and performance.
- Good feedback to training. Goal: More than average grade in class evaluation.

Year 2: Adaptive behavior of the system, providing more resources as needed/possible

Integration and Provisioning:

- Contribute and support integrated system to OSG VDT release. (Fermi)
- Contribute and support integrated system as a TeraGrid capability kit. (USC)
- Evaluate system on commercial cloud and one community (USC, Fermi)

Development:

- Management tools for multiple front-end plugins. (Fermi)
- Common interface to OSG and TeraGrid information and publishing services. (USC, Fermi)
- Common support for architecture dependent software installation and execution. (Fermi, USC)
- Develop techniques for reacting to failures of explicit requests too big to be satisfied as a unit (UCSD, USC)
- Publish statistics collected and improve the interpretation and initial automated response to error conditions (USC, UCSD)

Collaboration and Community: (ALL)

- Engage two new science/research communities who want to use both OSG and TeraGrid resources using the integrated system.
- Identify further aspects of integration of the original two systems based on the first year's experience.
- Survey existing and potential user communities for their needs.

Outreach and Training:

- Provide training modules for HPC University and OSG e-learning environments. (ALL)

Metrics:

- Testimonials from users acknowledging benefit from use of the software for scientific output: Goal: 3
- Number of jobs executing in parallel. Goal: > 30,000.
- Number of jobs per day. Goal: > 150,000
- Increase in number of user communities supported. Goal: +2.
- Use of cloud resource provisioning type. Goal: meet metric set in Year 1.
- Good feedback to training. Goal: More than "good" grade in class evaluations.

Year 3: Usability and Maintenance

Integration and Provisioning:

- Continue to deliver releases to VDT and TeraGrid and support the system for an expanding set of users. (Fermi, USC)

Development:

- Add factories for additional cloud (e.g. windows VMs) and distributed computing environments (e.g. Blue Water, ANL Leadership Class Facility). (Fermi, USC)
- Improve application and system monitoring capabilities, end-to-end. (Fermi, USC)

Collaboration and Community: (ALL)

- Engage two new science/research communities who want to use both OSG and TG resources using the integrated system.
- Identify further aspects of common extensions based on Year 2 experience.

Outreach and Training: (ALL)

- Deliver and enhance training modules for HPC University and OSG e-learning environments; and reach out to university classroom environments.

Metrics:

- Testimonials from users acknowledging benefit from use of the software for scientific output: Goal: 6 testimonials.
- Number of jobs executing in parallel. Goal: > 40,000.
- Number of jobs per day. Goal: > 200,000.
- Increase in number of user communities supported. Goal: +2.
- Use of cloud resource provisioning type. Goal: increase usage by 25% from Year 2.
- Good feedback to training. Goal: More than "good" grade in class evaluations.
- Community support for continuation of the software development and support.

C 7. Management plan

Dr. Deelman and Dr. Holzman will form the management team and oversee the overall proposed effort. Dr. Deelman and Dr. Holzman have extensive experience in collaborative efforts as described above. They will have weekly conference calls to discuss the technical program of work and progress. The project will hold face to face technical meetings at least semi-annually to ensure continued progress and synergy of the collaborative work. It is expected that the technical developers will be in contact on a continuing basis. The project will continue to engage the current user community as well as reach out to other potential users of the technology and other workflow tool developers. We will use Bugzilla, an open source bug tracking system to keep track of support issues. The software will be maintained in a common source code repository at sourceforge.net. Mailing lists and a wiki will be set up for both internal communications and communications with the user community.

The management team will meet regularly with the major user communities – SCEC, US CMS and Fermilab – as well as have "taking-stock" meetings quarterly to record requirements, progress and issues. The management team will meet on a quarterly basis with the Open Science Grid Executive Team and with the TeraGrid GIG/Science Gateway forum to ensure good communication and collaboration with the cyberinfrastructure service providers.

Collaborative Tools: Designs, technical implementation discussions and documentation will be captured on a collaborative wiki which is publicly available.

Planning and Progress monitoring: The activities and deliverables will be captured in an annual program of work (Work Breakdown Structure) that will be updated by the project team on a quarterly basis. Quarterly reports of work accomplished will be gathered from each member of the team and posted centrally.

Evaluation: We will measure the success of our project by the metrics defined in Section C6. Standard project management practices will be used to track progress and measure the status of the project.

References

[1] E. Seidel, "NSF and the National Cyberinfrastructure," Presentation at the Open Science Grid AHM, 2009.

[2] "Cyberinfrastructure Vision for 21st Century Discovery," National Science Foundation Cyberinfrastructure CouncilMarch 2007.

[3] "Open Science Grid." www.opensciencegrid.org

[4] "TeraGrid." http://www.teragrid.org/

[5] G. Juve, E. Deelman, K. Vahi, and G. Mehta, "Corral: A Resource Provisioning System for Scientific Workflow Applications," 2009 (in submission).

[6] I. Sfiligoi, "glideinWMS-A generic pilot-based Workload Management System," *Journal of Physics: Conference Series,* vol. 119, 2008.

[7] R. Adolphi and Others, "The CMS Experiment at the CERN LHC," *JINST 0803:S08004,* 2008.

[8] S. Belforte and others, "GlideCAF: A Late Binding Approach to the Grid," in *CHEP06*, 2006.

[9] R. Blair and others, "The CDF II Detector Technical Design Report " Fermilab preprint FERMILAB-Pub-96/390-E1996.

[10] S. Wojcicki and others, "The MINOS Collaboration " Fermilab NuMI-L-3371998.

[11] T. D. Collab.", "The D0 Upgrade: The Detector and its Physics," Fermilab preprint Fermilab Pub-96/357-E1996

[12] "Pegasus." http://pegasus.isi.edu

[13] E. Deelman, G. Singh, M.-H. Su, J. Blythe, Y. Gil, C. Kesselman, G. Mehta, K. Vahi, G. B. Berriman, J. Good, A. Laity, J. C. Jacob, and D. S. Katz, "Pegasus: a Framework for Mapping Complex Scientific Workflows onto Distributed Systems," *Scientific Programming Journal,* vol. 13, pp. 219-237, 2005.

[14] S. Callaghan, P. Maechling, E. Deelman, K. Vahi, G. Mehta, G. Juve, K. Milner, R. Graves, E. Field, D. Okaya, D. Gunter, K. Beattie, and T. Jordan, "Reducing Time-to-Solution Using Distributed High-Throughput Mega-Workflows - Experiences from SCEC CyberShake," in *Fourth IEEE International Conference on e-Science (e-Science 2008)* Indianapolis, Indiana, USA, 2008.

[15] E. Deelman, S. Callaghan, E. Field, H. Francoeur, R. Graves, N. Gupta, V. Gupta, T. H. Jordan, C. Kesselman, P. Maechling, J. Mehringer, G. Mehta, D. Okaya, K. Vahi, and L. Zhao, "Managing Large-Scale Workflow Execution from Resource Provisioning to Provenance Tracking: The CyberShake Example," *E-SCIENCE '06: Proceedings of the Second IEEE International Conference on e-Science and Grid Computing,* p. 14, 2006.

[16] G. Singh, M. Su, K. Vahi, E. Deelman, B. Berriman, J. Good, D. Katz, and G. Mehta, "Workflow task clustering for best effort systems with Pegasus," 2008.

[17] M. Litzkow, M. Livny, and M. Mutka, "Condor - A Hunter of Idle Workstations," in *Proc. 8th Intl Conf. on Distributed Computing Systems*, 1988, pp. 104-111.

[18] R. V. Budny, C. E. Kessel, J. E. Kinsey, F. Imbeaux, I. Voitsekhovitch, J. Candy, R. E. Waltz, T. Fujita, C. M. Greenfield, M. Murakami, A. C. C. Sips, W. A. Houlberg, and E. J. Doyle, "Transport Physics of Hybrid Scenario Plasmas in the International Multi-Tokamak Database and Implications for ITER," in *Proc of the 21st IAEA Fusion Energy Conference*, 2006.

[19] "The Virtual Data Toolkit," http://vdt.cs.wisc.edu/, 2006.

[20] J. M. Wing, "Computational thinking," *Communications of the ACM,* vol. 49, pp. 33-35, 2006.

[21] G. Singh, C. Kesselman, and E. Deelman, "Optimizing Grid-Based Workflow Execution," *Journal of Grid Computing,* vol. 3, pp. 201-219, 2005.

[22] G. Singh, C. Kesselman, and E. Deelman, "Performance Impact of Resource Provisioning on Workflows," USC ISI Technical Report2005.

[23] G. B. Berriman, E. Deelman, J. Good, J. Jacob, D. S. Katz, C. Kesselman, A. Laity, T. A. Prince, G. Singh, and M.-H. Su, "Montage: A Grid Enabled Engine for Delivering Custom Science-Grade Mosaics On Demand," in *SPIE Conference 5487: Astronomical Telescopes*, 2004.

[24] T. Oinn, P. Li, D. B. Kell, C. Goble, A. Goderis, M. Greenwood, D. Hull, R. Stevens, D. Turi, and J. Zhao, "Taverna/myGrid: Aligning a Workflow System with the Life Sciences Community," in *Workflows in e-Science*, I. Taylor, E. Deelman, D. Gannon, and M. Shields, Eds.: Springer, 2006.

[25] A. Damjanovi and others, "Open Science Grid study of the coupling between conformation and water content in the interior of a protein," *Journal of Physical Chemistry, B,* 2008.

[26] M. Jones, B. Ludascher, D. Pennington, and A. Rajasekar, "Data Integration and Workflow Solutions for Ecology," in *Data Integration in Life Sciences*, 2005.

[27] P. Maechling, E. Deelman, L. Zhao, R. Graves, G. Mehta, N. Gupta, J. Mehringer, C. Kesselman, S. Callaghan, D. Okaya, H. Francoeur, V. Gupta, Y. Cui, K. Vahi, T. Jordan, and E. Field, "SCEC CyberShake Workflows---Automating Probabilistic Seismic Hazard Analysis Calculations," in *Workflows for e-Science*, I. Taylor, E. Deelman, D. Gannon, and M. Shields, Eds.: Springer, 2006.

[28] J. Muench, H. Francoeur, D. Okaya, Y. Cui, P. Maechling, E. Deelman, G. Mehta, R. Moore, and T. Jordan, "SCEC Earthworks Science Gateway: Widening SCEC Community Access to the TeraGrid," TeraGrid 2006 Conference, 2006.

[29] D. A. Brown, P. R. Brady, A. Dietz, J. Cao, B. Johnson, and J. McNabb, "A Case Study on the Use of Workflow Technologies for Scientific Analysis: Gravitational Wave Data Analysis," in *Workflows for e-Science*, I. Taylor, E. Deelman, D. Gannon, and M. Shields, Eds.: Springer, 2006.

[30] B. Abbott, A. Baranovski, M. Diesburg, G. Garzoglio, T. Kurca, and P. Mhashilkar, "DZero Data-Intensive Computing on the Open Science Grid " *Journal of Physics: Conference Series,* vol. 119 2008.

[31] P. C. Bhat, "Run II physics at the Fermilab Tevatron and advanced analysis methods Nuclear Instruments and Methods," *Physics Research Section A,* vol. 502, 2003.

[32] G. L. Bayatyan, M. D. Negra, H. A. Foà, and A. Petrilli, "CMS computing : Technical Design Report," CERN Report Number CERN-LHCC-2005-022005

[33] "Atlas Collaboration, Atlas - Technical Proposal," CERN Report Number CERN-LHCC94-431994

[34] "OSG Grid Schools." https://twiki.grid.iu.edu/bin/view/Education/GridWorkshops

[35] C. Hoffa, G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman, and J. Good, "On the Use of Cloud Computing for Scientific Workflows," 2008, pp. 640-645.

[36] B. Berriman, A. Bergou, E. Deelman, J. Good, J. Jacob, D. Katz, C. Kesselman, A. Laity, G. Singh, M.-H. Su, and R. Williams, "Montage: A Grid-Enabled Image Mosaic Service for the NVO," in *Astronomical Data Analysis Software & Systems (ADASS) XIII*, 2003.

[37] S. Callaghan, P. Maechling, E. Deelman, K. Vahi, G. Mehta, G. Juve, K. Milner, R. Graves, E. Field, D. Okaya, and T. Jordan, "Reducing Time-to-Solution Using Distributed High-Throughput Mega-Workflows: Experiences from SCEC CyberShake," in *4th IEEE International Conference on e-Science (e-science'08)*: (submitted), 2008.

[38] *Workflows in e-Science*. I. Taylor, E. Deelman, D. Gannon, and M. Shields, Eds.: Springer, 2006.

[39] G.Bagliesi and others, "Debugging Data Transfers in CMS," in *CHEP09*, 2009.

[40] D. Bradley and others, "Use of glide-ins in CMS for production and analysis," in *CHEP09*, 2009.

[41] D. Bradley and others, "Interoperability and scalability within glideinWMS," in *CHEP09*, 2009.

[42] J.Andreeva and others, "CMS Analysis Operations," in *CHEP09*, 2009.

[43] I. Sfiligoi, D. Bradley, and M. Livny, "Pseudo-interactive monitoring in distributed computing," in *CHEP09*, 2009.

[44] I. Sfiligoi and others, "The pilot way to Grid resources using glideinWMS," in *CSIE09*, 2009.

[45] A. Fanfani, "Commissioning Distributed Analysis at the CMS Tier-2 Centers," in *CHEP09*, 2009.

[46] "condor_glidein." http://www.cs.wisc.edu/condor/glidein.
[47] B. Sotomayor and L. Childers, *Globus toolkit 4 programming Java services*: Elsevier; Morgan Kaufmann, 2006.
[48] V. Welch, F. Siebenlist, I. Foster, J. Bresnahan, K. Czajkowski, J. Gawor, C. Kesselman, S. Meder, L. Pearlman, and S. Tuecke, "Security for Grid Services.," in *Twelfth International Symposium on High Performance Distributed Computing (HPDC-12)*, 2003.
[49] W. Allcock, A. Chervenak, I. Foster, L. Pearlman, V. Welch, and M. Wilde, "Globus Toolkit Support for Distributed Data-Intensive Science," in *International Conference on Computing in High Energy and Nuclear Physics*, 2001.
[50] A. Chervenak, E. Deelman, I. Foster, L. Guy, W. Hoschek, A. Iamnitchi, C. Kesselman, P. Kunst, M. Ripenu, B. Schwartzkopf, H. Stockinger, K. Stockinger, and B. Tierney, "Giggle: A Framework for Constructing Sclable Replica Location Services.," in *Proceedings of Supercomputing 2002 (SC2002)*, 2002.
[51] Amazon.com, "Elastic Compute Cloud (EC2)." http://aws.amazon.com/ec2
[52] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L. Youseff, and D. Zagorodnov, "The Eucalyptus Open-source Cloud-computing System," in *Cloud Computing and its Applications*, 2008.
[53] "Nimbus Science Cloud." http://workspace.globus.org/clouds/nimbus.html
[54] M. Livny, "Matchmaking: Distributed Resource Management for High Throughput Computing," in *IEEE International Symposium on High Performance Distributed Computing*, 1998.

Ewa Deelman

Information Sciences Institute
University of Southern California
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90292-6695

310 488-8408
310 823-6714
deelman@isi.edu
http://www.isi.edu/~deelman

Education

Rensselaer Polytechnic Institute, Troy, NY	Ph.D. Computer Science	1997
SUNY New Paltz, NY	M.S. Computer Science	1991
Wells College, Aurora, New York	B.A. Mathematics	1987

Appointments

2007	Project Leader, Information Sciences Institute
2003	Research Assistant Professor, Computer Science Department, USC
2002	Research Team Leader, Information Sciences Institute, USC
2000	Computer Scientist, Information Sciences Institute, USC
1997	Sr. Software Developer, University of California, Los Angeles

Related Publications

[1] Singh, G., C. Kesselman, and E. Deelman, An End-to-End Framework for Provisioning-Based Resource and Application Management, *IEEE Systems Journal, (3) 1, March 2009*

[2] Callaghan, S., P. Maechling, E. Deelman, et al., "Reducing Time-to-Solution Using Distributed High-Throughput Mega-Workflows - Experiences from SCEC CyberShake", Fourth IEEE International Conference on e-Science (e-Science 2008), Indianapolis, Indiana, USA, 2008

[3] Juve, G., E. Deelman, "Resource Provisioning Options for Large-Scale Scientific Workflows", *Third International Workshop on Scientific Workflows and Business Workflow Standards in e-Science (SWBES)* in conjunction with Fourth IEEE International Conference on e-Science (e-Science 2008), 2008

[4] Deelman, E., G. Singh, M. Livny, B. Berriman, J. Good, "The Cost of Doing Science on the Cloud: The Montage Example" Proceeding of Super Computing 2008, Austin , Texas, November 2008

[5] Gil, Y. and E. Deelman, "Examining the Challenges of Scientific Workflows", IEEE Computer, 40(12), p 24-32, 2007

Other Significant Publications

[1] Deelman, E., D. Gannon, M. Shields, I. Taylor, "Workflows and e-Science: An overview of workflow system features and capabilities", *Future Generations of Computer Systems*, July 2008.

[2] Miles, S. P. Groth, E. Deelman, K. Vahi, G. Mehta, and L. Moreau. "Provenance: The bridge between experiments and data". *Computing in Science and Engineering*, 2008.

[3] Taylor, I., E. Deelman, D. Gannon, and M. Shields (*Eds*), *Workflows in e-Science*: Springer, 2006.

[4] Deelman, E. et al., "Managing Large-Scale Workflow Execution from Resource Provisioning to Provenance tracking: The CyberShake Example," in e-Science, Amsterdam, The Netherlands, 2006. (best paper award)

[5] Deelman, E. , et al., "What Makes Workflows Work in an Opportunistic Environment?," Concurrency and Computation: Practice and Experience, vol. 18, 2005.

Awards
Best Paper award, 2nd e-Science Conference, Amsterdam 2006. Best Paper award, 15th Workshop on Parallel and Distributed Simulation, Lake Arrowhead, CA, USA; 15-18 May 2001.General Electric Foundation Fellowship , Rensselaer Polytechnic Institute, 1993-1994. Special Distinction in the field of Mathematics, Wells College, 1987.

Synergistic Activities
Research tools: Developed software that have seen great use in the research community including Pegasus (Planning for Execution in Grids, pegasus.isi.edu), and the Metadata Catalog Service (MCS, www.isi.edu/~deelman/MCS).
Mentoring: Providing research internships for undergraduate women in Computer Science as part of the CRA-W program.
Service:

- Associate Editor responsible for Grid Computing for the Scientific Programming Journal
- Chair: Workshop on Workflows in Support of Large-Scale Science, in Conjunction with HPDC, June 2006.
- Co-Chair: Workshop on Workflows in Support of Large-Scale Science, in Conjunction with HPDC, June 2007.
- Co-Chair: The Third Grid Applications and Middleware Workshop, in Conjunction with PPAM 2007
- Chair: Global Grid Forum Workflow Management Research Group
- Member of USC Women in Science and Engineering (WiSE)

Program committee Member (2009- 2005): 4th Int'l Workshop on Workflow systems in e-Science (WSES09), High Performance Distributed Computing (HPDC 2009), SC 2008, Grid 2008, High Performance Distributed Computing (HPDC 2008), e-Science 2007, Bangalore, India, December 2007, High Performance Distributed Computing (HPDC 2007), IEEE International Conference on Grid Computing (Grid 2007) (Vice Program chair for Scheduling, Resource Management and Runtime Environments), Second International Conference on Grid computing, high-performance and Distributed Applications" (GADA'07), Workflow Systems in e-Science 2007, International Conference on Computational Science (ICCS2007) , IEEE International Symposium on Cluster Computing and the Grid, CCGrid 2007,Fourth International Symposium on Parallel and Processing and Applications (ISPA'06), (Software and Applications Track), Sorrento, Italy, December 2006., "Scientific workflow management in e-Science", in conjunction with the International Conference on Computational Science (ICCS2006), May 2006., IEEE/ACM CCGrid conference 2006, SSDBM 2006, ICIW 2006 Web Service-based Systems and Applications track, February 2006, International Workshop on Scientific Instruments and Sensors on the Grid, December 2005, Applications Grid Workshop, September 2005, IEEE/ACM International Workshop on Grid Computing (Grid 2005), SSDBM 2005, IEEE/ACM CCGrid conference 2005, International Workshop on Grid and Peer-to-Peer based Workflows, 2005, Semantic Infrastructure for Grid Applications at CCGrid 2005, European Grid Conference 2005, Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA), held in conjunction with the International Conference on Parallel Processing (ICPP-2005)

Collaborators Matthew Arrott (UCSD), Bruce Berriman (Caltech), Ann Chervenak (USC), Kent Blackburn (Caltech), Duncan Brown (Syracuse University), Mark Ellisman (UCSD), Dennis Gannon (Indiana University), Ian Foster (University of Chicago), Yolanda Gil (USC), Jeffrey Grethe (UCSD), Carole Goble (University of Manchester), John Good (Caltech), Mary Hall (USC), Robert Hanisch (John Hopkins), Joseph Jacobs (JPL), Thomas Jordan (USC), Daniel Katz (LSU), Carl Kesselman (USC), Scott Koranda (UWM), Philip Maechling (USC), Miron Livny (U. of Wisc.), David Meyers (USC), Simon Miles (King's College London), Luc Moreau (Southampton University), Jim Myers (NCSA), Jarek Nabrzyski (Poznan Supercomputing Center), Veronika Nefedova (ANL), Steven Peltier (UCSD), Rizos Sakellariou (University of Manchester), Joel Saltz (OSU), Matthew Shields (Cardiff University), Boleslaw Szymanski (RPI), Ian Taylor (Cardiff University), Valery Taylor (UTexas), Michael Wilde (ANL), Robert Wilhelmson (NCSA), Roy Williams (Caltech)

PhD Thesis Advisor: Prof. Boleslaw Szymanski, Rensselaer Polytechnic Institute
PhD Students: Gurmeet Singh, Gideon Juve

Burt Holzman

Fermi National Accelerator Laboratory
M/S 234
P.O. Box 500
Batavia, IL 60510

630 840 5753

burt@fnal.gov

Professional Preparation

Carnegie Mellon University	Mechanical Engineering	B.S. 1995
University of Illinois at Chicago	Physics	M.S. 1997
University of Illinois at Chicago	Physics	Ph.D. 2000
Brookhaven National Laboratory	Nuclear Physics	2001-2003

Appointments

2008	Associate Department Head, CMS Tier 1, Fermi National Accelerator Laboratory
2007	Group Leader, CMS Grid Services, Fermi National Accelerator Laboratory
2006	Project Lead, GlideinWMS, Fermi National Accelerator Laboratory
2005	Computing Professional, Fermi National Accelerator Laboratory
2004	Assistant Scientist, Brookhaven National Laboratory
2001	Postdoctoral Researcher, Brookhaven National Laboratory

Related Publications

[1] The glideinWMS software system: http://www.uscms.org/~glideinWMS

[2] I. Sfiligoi, D. Bradley, B. Holzman, P. Mhashilkar, S. Padhi, F. Würthwein, "The pilot way to Grid resources using glideinWMS", Proceedings of CSIE 2009: 428-432 (accepted for publication, 2009)

[3] D. Bradley *et al*., "Use of glide-ins in CMS for production and analysis", J. Phys. Conf. Ser. (accepted for publication, 2009)

[4] I. Sfiligoi and B. Holzman, "An objective comparison test of workload management systems", J. Phys. Conf. Ser. **119:** 062043 (2007)

Other Significant Publications

[1] B. Back *et al.*, "The PHOBOS perspective on discoveries at RHIC", Nucl. Phys. A757: 28-101 (2005)

[2] B. Back *et al.*, "Charged particle multiplicity near mid-rapidity in central Au+Au collisions at $\sqrt{s}$ = 56 AGeV and 130 AGeV", Phys. Rev. Lett. 85: 3100-3104 (2000)

[3] B. Back *et al.*, "Centrality dependence of charged hadron transverse momentum spectra in d+Au collisions at $\sqrt{s_{NN}}$ = 200 GeV", Phys. Rev. Lett. 91: 072302 (2003)

[4] B. Back *et al.*, "The significance of the fragmentation region in ultrarelativistic heavy ion collisions", Phys. Rev. Lett. 91: 052303 (2003)

[5] B. Back *et al.,* "Transverse momentum and rapidity dependence of HBT correlations in Au+Au collisions at $\sqrt{s_{NN}}$ = 62.4 GeV and 200 GeV", Phys. Rev. C73: 031901 (2006)

Awards

University Fellowship, University of Illinois at Chicago, 1997-1998

Synergistic Activities

Research tools: Developed and refined software projects in wide use by the Open Science Grid Community, including glideinWMS (www.uscms.org/~glideinWMS) and the OSG Information Services project (twiki.grid.iu.edu/bin/view/InformationServices).

Collaborators

- Collaborators and Co-Editors:

Balamurali Ananthan (Tech-X), Emanouil Atanassov (BAS), Burak Alver (MIT), Birger Back (ANL), Mark Baker (BNL), Maarten Ballintijn (MIT), Don Barton (BNL), Russell Betts (IIT), Richard Bindel (UMD), Brian Bockelman (UNL), Daniel Bradley (UW), Wit Buzsa (MIT), Zhengwei Chai (MIT), Vasundhara Chetluru (UIC), Edmundo Garcia (UIC), Gabriele Garzoglio (FNAL), Tomasz Gburek (INP), Kristian Gulbrandsen (MIT), Ian Harnarine (UIC), Conor Henderson (MIT), Aneta Iordanova (UIC), Jay Kane (MIT), Piotr Kulinich (MIT), Chia-Ming Kuo (NCU), Wei Li (MIT), Willis Lin (NCU), Constantin Loizides (MIT), Steve Manly (Rochester), Alice Mignerey (UMD), Rachid Nouicer (BNL), Andrzej Olszewski (INP), Sanjay Padhi (UCSD), Robert Pak (BNL), Haifeng Pi (UCSD), Corey Reed (MIT), Eric Richardson (UMD), Christof Roland (MIT), Gunther Roland (MIT), Joe Sagerer (UIC), Iouri Sedykh (BNL), Igor Sfiligoi (UCSD), Chadd Smith (UIC), Maciej Stankiewicz (BNL), Peter Steinberg (BNL), George Stephans (MIT), Andrei Sukhanov (BNL), Artur Szostak (BNL), Nicholas Thackray (CERN), Anthony Tiradani (FNAL), Marguerite Tonjes (UMD), Steve Traylen (CERN), Adam Trzupek (INP), Andreas Unterkircher (CERN), Gerrit van Nieuwenhuizen (MIT), Sergei Vaurynovich (MIT), Robin Verdier (MIT), Gabor Veres (MIT), Pete Walters (Rochester), John Weigand (FNAL), Ed Wenger (MIT), Donald Willhelm (UMD), Frank Wolfs (Rochester), Barbara Wosiek (INP), Krzystof Wozniak (INP), Frank Würthwein (UCSD), Shaun Wyngaardt (BNL), Bolek Wyslouch (MIT)

Graduate Advisor: R. Russell Betts (IIT)
Postdoctoral Sponsors: Mark Baker (BNL), Peter Steinberg (BNL)
Students: Chia-Ming Kuo (NCU)

Frank Würthwein
Professor of Physics, University of California, San Diego

Professional Expertise

HEP: Frank Würthwein's research focuses on Experimental Particle Physics with the CMS experiments. In preparation for CMS data taking, Würthwein's group is now focusing on di-lepton physics studying WW production, as well as search for new physics in di-leptons. Prior to this, Würthwein's group discovered WZ production, made the first measurement of ZZ production at the Tevatron, and lead the Higgs search in the WW final state at CDF.

CMS Computing: In the context of CMS, Würthwein's group is operating a CMS tier-2 center, which is part of the Data Intensive Science University Network (DISUN) for which Würthwein is the technical lead. In addition, Würthwein leads the CMS Distributed Computing Tools (DCT) group, which is responsible for CMS user analysis computing as well as Monte Carlo Production on the Open Science Grid. Würthwein is moreover co-leading commissioning of the global CMS computing infrastructure.

Open Science Grid: As Chair of the Interim Executive Board of the OSG (12/2004 – 1/2006) Würthwein oversaw the formation of the OSG Consortium and Project, as well as the deployment of the first OSG Production Grid Infrastructure. He was thus leading the OSG through its governing, technical, and deployment processes towards its present form. With the transition to its present governance structure, Würthwein has taken on the role of Application Coordinator of the OSG together with Torre Wenaus (BNL). He is also responsible for the work area "scalability, reliability, and usability" in the OSG.

Education & Employment

UC San Diego	Associate Professor (Physics), 2003-Present
Fermi National Laboratory	Guest Scientist in computing division, 2002-2004
MIT	Assistant Professor (Physics), 1999-2003
Post-doctoral Institution:	Caltech, Millikan Fellow (Physics), 1995-1998
	Caltech, Senior Research Associate, 1998-1999
Graduate Institution:	Cornell University, Ph.D. (Physics), 1995
Undergraduate Institution:	Universität Heidelberg, Vordiplom (Physics), 1988

Selected Professional Positions

Co-leader of Computing Commissioning in CMS, since 1/2007
Applications Coordinator for the Open Science Grid, since 1/2006
Technical Lead of DISUN, since 6/2005
Lead of US CMS Distributed Computing Tools, since 6/2005
Chair of Interim Executive Board of the Open Science Grid, 12/2004-1/2006.
Assoc. Head of Computing Infrastructure for the CDF experiment, 2002-2004.
Co-leader of the Offline Project for the CDF Experiment, 2003-2004

Selected Publications

Observation of WZ Production, PRL 98 (2007) 161801
Strong Evidence for ZZ Production, PRL 100 (2008) 201801
Search for Higgs Boson decaying to two W bosons at CDF, Phys. Rev. Lett., 102, 021802 (2009)
"Use of glide-ins in CMS production and analysis", Proc. Computing in High Energy Physics (2009)
"Debugging Data Transfers in CMS", Proc. Computing in High Energy Physics (2009)
"Science on the Grid with CMS at the LHC", SciDAC 2008

Gabriele Garzoglio

Fermi National Accelerator Laboratory
M/S 120
P.O. Box 500
Batavia, IL 60510

630 840 6470

garzoglio@fnal.gov

Professional Preparation

University of Genova, Italy	Laurea in Physics	M.S.-equivalent 1996
University of DePaul, Chicago	Computer Science	Ph.D. 2006

Appointments

2006	Project Lead, VO Services project, Fermi National Accelerator Laboratory
2005	Group Lead, Open Science Grid group, Fermi National Accelerator Laboratory
2005	Project Lead, Resource Selection project, Fermi National Accelerator Laboratory
2001	Computing Professional, Computing Division, Fermi National Accelerator Laboratory
1998	Guest Engineer, Beam Division, Fermi National Accelerator Laboratory
1996	Guest Scientist, Particle Physics Division, Fermi National Accelerator Laboratory

Selected Related Publications

[1] Gabriele Garzoglio et al., "Definition and Implementation of a SAML-XACML Profile for Authorization Interoperability Across Grid Middleware in OSG and EGEE", Journal of Cluster Computing (online) on Apr 2009, DOI: 10.1007/s10723-009-9117-4

[2] Adriana Iamnitchi, Shyamala Doraimani, Gabriele Garzoglio, "Workload Characterization in a High-Energy Data Grid and Impact on Resource Management", Journal of Cluster Computing on Jan 23, 2009. DOI: 10.1007/s10586-009-0081-3

[3] B. Abbott, A. Baranovski, M. Diesburg, G. Garzoglio, T. Kurca, P. Mhashilkar, "DZero Data-Intensive Computing on the Open Science Grid", Journal of Physics: Conference Series (JPCS), Sep 2008

[4] A. Iamnitchi, S. Doraimani, G. Garzoglio, "Filecules in High-Energy Physics: Characteristics and Impact on Resource Management", Proceedings of the 15th IEEE International Symposium on High Performance Distributed Computing (HPDC-15), Paris, France, June 2006

[5] G. Garzoglio, T. Levshina, P. Mhashilkar, S. Timm, "ReSS: a Resource Selection Service for the Open Science Grid.", Proceedings of the International Symposium of Grid Computing (ISGC07), March 2007, Taipei, Taiwan

[6] G. Garzoglio, A.Baranovski, H. Koutaniemi, L. Lueking, S. Patil, R. Pordes, A. Rana, I. Terekhov, S. Veseli, J. Yu, R. Walker, V. White, "The SAM-GRID project: architecture and plan." Nuclear Instruments and Methods in Physics Research, Section A, NIMA14225, vol. 502/2-3 pp 423 - 425

Other Significant Publications

[1] The E835 Collaboration (M Andreotti et al.), "Experiment E835 at Fermilab"
Published in Nuclear Inst. and Methods in Physics Research, A, NIM A Vol 519/3 pp 558-609

[2] The E835 Collaboration (M. Andreotti et al.) "Precision measurements of the total and partial widths of the psi(2S) charmonium meson with a new complementary-scan technique in antiproton-proton annihilations", Physics Letters B 654 (2007) 74-79

[3] By The E835 Collaboration (M. Andreotti et al.) "Measurement of the branching ratios psi' → e^+e^-, psi' → *J*/psi pi pi, and psi' → *J*/psi eta " Phys. Rev. D 71, 032006 (2005)

[4] By The E835 Collaboration (M. Andreotti et al.) "Measurement of the angular distribution in anti-p p → psi(2S) → e+ e-" Physics Letters B Volume 610, Issues 3-4, 24 March 2005, Pages 177-182

Synergistic Activities

Research tools: Managed and developed software projects in use by the Open Science Grid Community, including the Resource Selection Service Project (https://twiki.grid.iu.edu/bin/view/ResourceSelection) and the VO Services Infrastructure (http://www.fnal.gov/docs/products/voprivilege), and in use by the

Worldwide LHC Computing Grid, in particular the Authorization Interoperability Project (http://www.fnal.gov/docs/products/voprivilege/focus/AuthZInterop/info.html).

Collaborators

- Collaborators and Co-Editors:

Brad Abbott (OU), Ian Alderman (UWisc), Mine Altunay (FNAL), Rachana Ananthakrishnan (ANL), Andrew Baranovski (FNAL), Joe Bester (ANL), Keith Chadwick (FNAL), Vincenzo Ciaschini (INFN), Yuri Demchenko (UVA), Mike Diesburg (FNAL), Andrea Ferraro (INFN), Alberto Forti (INFN), David Groep (NIKHEF), Ted Hesselroth (FNAL), John Hover (BNL), Adriana Iamnitchi (USF), Oscar Koeroo (NIKHEF), Chad La Joie (SWITCH), Lee Lueking (FNAL), Tibor Kurca (IPNL), Tanya Levshina (FNAL), Parag Mhashilkar (FNAL), Zach Miller (UWisc), Jay Packard (BNL), Siddharth Patil (UTA), Ljubomir Perkovic (DePaul), Ruth Pordes (FNAL), Abhishek Rana (UCSD), , Alain Roy (UWisc), Håkon Sagehaug (BCCS), Igor Sfiligoi (FNAL), Neha Sharma (FNAL), Frank Siebenlist (ANL), Rick St. Denis (UGlasgow), Todd Tannenbaum (UWisc), Igor Terekhov (FNAL), Valerio Venturi (INFN), Sinisa Veseli (FNAL), Jae Yu (UTA), Rod Walker (IC), John Weigand (FNAL), Vicky White (FNAL), . Frank Wuerthwein (UCSD)

Graduate Advisor: Ljubomir Perkovic (DePaul)
Students: Bimal Balan (UTA), Sankalp Jain (UTA), Hannu Koutaniemi (Espoo-Vanta IT), Parag Mhashilkar (UTA), Vijay Murthi (UTA), Aditya Nishandar (UTA), Siddharth Patil (UTA), Anoop Rajendra (UTA), Abhishek Rana (UTA), Sudhamsh Reddy (UTA)

SUMMARY PROPOSAL BUDGET

YEAR 1

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Southern California

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Ewa Deelman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Ewa Deelman - PI/Research Asst Prof**	1.20	0.00	0.00	$ **15,940**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	1.20	0.00	0.00	**15,940**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	9.60	0.00	0.00	**73,938**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**89,878**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**26,963**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**116,841**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**6,731**	
2. FOREIGN				**6,781**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**2,622**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**8,629**	
5. SUBAWARDS				**276,420**	
6. OTHER				**27,862**	
TOTAL OTHER DIRECT COSTS				**315,533**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**445,886**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **F&A (Rate: 26.0000, Base: 219466) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**89,981**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**535,867**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **535,867**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Ewa Deelman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 1

**** I- Indirect Costs**
ISI Facilities (Rate: 15.0000, Base 219466)

SUMMARY PROPOSAL BUDGET

YEAR 2

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Southern California

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Ewa Deelman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Ewa Deelman - PI/Research Asst Prof**	1.20	0.00	0.00	$ **16,743**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	1.20	0.00	0.00	**16,743**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	9.60	0.00	0.00	**77,662**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**94,405**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**28,321**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**122,726**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**7,070**	
2. FOREIGN				**7,123**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**2,622**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**8,629**	
5. SUBAWARDS				**279,093**	
6. OTHER				**29,266**	
TOTAL OTHER DIRECT COSTS				**319,610**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**456,529**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **F&A (Rate: 26.0000, Base: 177437) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**72,750**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**529,279**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **529,279**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Ewa Deelman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 2

**** I- Indirect Costs**
ISI Facilities (Rate: 15.0000, Base 177437)

SUMMARY PROPOSAL BUDGET

YEAR 3

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Southern California

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Ewa Deelman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Ewa Deelman - PI/Research Asst Prof**	1.20	0.00	0.00	$ **17,578**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	1.20	0.00	0.00	**17,578**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	9.60	0.00	0.00	**81,116**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**98,694**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**29,608**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**128,302**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**7,423**	
2. FOREIGN				**7,478**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**2,622**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**8,591**	
5. SUBAWARDS				**285,025**	
6. OTHER				**30,596**	
TOTAL OTHER DIRECT COSTS				**326,834**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**470,037**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **F&A (Rate: 26.0000, Base: 185012) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**75,855**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**545,892**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **545,892**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Ewa Deelman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 3

**** I- Indirect Costs**
ISI Facilities (Rate: 15.0000, Base 185012)

SUMMARY PROPOSAL BUDGET

Cumulative

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of Southern California

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Ewa Deelman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Ewa Deelman - PI/Research Asst Prof**	3.60	0.00	0.00	$ **50,261**	$
2.					
3.					
4.					
5.					
6. ( ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	3.60	0.00	0.00	**50,261**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**3**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	28.80	0.00	0.00	**232,716**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**282,977**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**84,892**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**367,869**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**21,224**	
2. FOREIGN				**21,382**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**7,866**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**25,849**	
5. SUBAWARDS				**840,538**	
6. OTHER				**87,724**	
TOTAL OTHER DIRECT COSTS				**961,977**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**1,372,452**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)					
TOTAL INDIRECT COSTS (F&A)				**238,586**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**1,611,038**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **1,611,038**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Ewa Deelman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

C ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Budget Justification USC

The proposed work involves two pieces of software: Corral, developed at the USC Information Sciences Institute (ISI), and glideinWMS, developed at Fermi National Accelerator Laboratory (FNAL). We propose to develop a robust, integrated system composed of the two pieces and provide the capabilities described in the proposal and tasks outlined in Section C6. In total we have allocated the effort of approximately two software developers to the overall project, one developer at ISI and one at FNAL. Most of the effort on the part of Dr. Deelman is geared toward technical and overall project leadership. Dr. Holzman will also contribute to project leadership and will coordinate the effort at FNAL. Dr. Würthwein will lead the work at UCSD. PIs Deelman and Holzman will work closely together to lead the development of the proposed system and make sure that the two development teams have common goals, jointly develop system design documents, use a common code versioning system, etc.

Holzman (co-PI) and Garzoglio (Sr. Personnel) from FNAL will participate in the project at no cost. Their time will be contributed by the CMS project and the Computing Division at FNAL. CMS will fund one month per year of Holzman's effort and the Computing Division will fund two weeks per year of Garzoglio's effort. He will provide technical expertise in the area of Grid services.

UCSD will support the testing and deployment of the integrated system on the Open Science Grid (OSG). UCSD's effort will include automated error recovery and failure analysis. Additionally, we will rely on the expertise and participation of Igor Sfiligoi, who is moving from FNAL to UCSD and who is the lead architect of the glideinWMS. Sfiligoi will contribute to the design of CorralWMS and work with the Graduate Research Associate at UCSD.

Each institution allocated a travel budget to support in-person meetings that will enhance the interactions between the projects participants (in addition to conference calls, mailing lists, online chats, and in person meetings at other project meetings, such as for example OSG All Hands Meeting).

All team members will contribute to community engagement effort to reach out to new communities and assure relevance of the new developments for the users.
This proposal meets the challenge set out by the President and Congress to stimulate the economy through the creation of jobs. If funded, this project will contribute to the development of economical innovation and future economic security through the involvement of two professional software developers in conjunction with one graduate student. Consequently, the long term impact will be the generation of new knowledge and possible development of new ideas and technologies.

Staff Salaries (ITEM A) – This is derived by summing for the identified staff member(s) their (Monthly salary * Monthly Level of Effort) * Escalation Staff Salary)

Fringe Benefits (ITEM C) – This is derived by applying the Fringe Benefit rate to the Fringe salary base. The Fringe salary base consists of Staff and Faculty salaries.

Common (ITEM G.6) –This is the cost of ISI's executive office and is derived by applying the Common Rate to the Common base. The Common base consists of ISI staff salary, faculty salary, GRA stipends and student wages.

Computing Services (ITEM G.4) – This covers the cost of ISI's Information Processing Center. This is charged at a flat rate per month times the level of effort per month for staff, faculty and GRAs and students.

Facilities and Administration (ITEM I) – This is derived by multiplying the Modified Total Direct Cost (MTDC) per month by the Facilities and Administration rate. (Formerly termed Indirect Cost)

ISI Facilities (ITEM I) – This covers the cost of ISI's rent, power, and office equipment, calculated as a percentage of MTDC. This is derived by multiplying the Modified Total Direct Cost (MTDC) per month by the ISI Facilities rate.

Materials & Supplies (ITEM G.1) – M&S costs include the following:
$219 ($200 + tax) per month for miscellaneous materials & supplies.

Travel (ITEM E.1) – This covers the cost of an annual trip for 2 people, 3 days to Washington DC for Sponsor Meetings, an annual trip for 2 people, 3 days to Chicago, IL (Fermi Lab) for technical meetings, and an annual trip for 2 people, 3 days to San Diego, CA (UCSD) for technical meetings.

Travel (ITEM E.2) – This covers the cost of an annual conference trip for 2 people, 3 days to London, England.

Subcontractors (ITEM G.5) – The following are the participating subcontractors:

Fermi National Accelerator Laboratory
University of California – San Diego

Please see their individual subcontract budgets for their breakdown of costs & budget justification.

SUMMARY PROPOSAL BUDGET

YEAR 1

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
Fermi Research Alliance, LLC

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Burt Holzman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Burt Holzman - Computing Serv Speciailist**	0.00	0.00	0.00	$ **0**	$
2. **Gabrielle Garzoglio - Applic Dev & Sys Analyst**	0.00	0.00	0.00	**0**	
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**2**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.00	**0**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	11.98	0.00	0.00	**68,010**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**68,010**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**38,446**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**106,456**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**10,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**0**	
TOTAL OTHER DIRECT COSTS				**0**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**116,456**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **Equipment & Materials & Supplies (Rate: 16.0300, Base: 0) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**83,544**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**200,000**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **200,000**	$
M. COST SHARING PROPOSED LEVEL $ **0**	AGREED LEVEL IF DIFFERENT $				

PI/PD NAME
Burt Holzman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 1

**** I- Indirect Costs**
Salaries (Rate: 77.4900, Base 106457)
Travel (Rate: 10.5000, Base 10000)

SUMMARY PROPOSAL BUDGET

YEAR 2

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
Fermi Research Alliance, LLC

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Burt Holzman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Burt Holzman - Computing Serv Speciailist**	0.00	0.00	0.00	$ **0**	$
2. **Gabrielle Garzoglio - Applic Dev & Sys Analyst**	0.00	0.00	0.00	**0**	
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**2**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.00	**0**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	11.52	0.00	0.00	**68,010**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**68,010**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**38,446**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**106,456**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**10,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**0**	
TOTAL OTHER DIRECT COSTS				**0**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**116,456**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **Equipment & Materials & Supplies (Rate: 16.0300, Base: 0) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**83,544**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**200,000**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **200,000**	$
M. COST SHARING PROPOSED LEVEL $ **0**	AGREED LEVEL IF DIFFERENT $				

PI/PD NAME
Burt Holzman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 2

**** I- Indirect Costs**
Salaries (Rate: 77.4900, Base 106457)
Travel (Rate: 10.5000, Base 10000)

SUMMARY PROPOSAL BUDGET

YEAR 3

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
Fermi Research Alliance, LLC

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Burt Holzman

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Burt Holzman - Computing Serv Speciailist**	0.00	0.00	0.00	$ **0**	$
2. **Gabrielle Garzoglio - Applic Dev & Sys Analyst**	0.00	0.00	0.00	**0**	
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**2**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.00	**0**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	11.08	0.00	0.00	**68,010**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**68,010**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**38,446**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**106,456**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**10,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**0**	
TOTAL OTHER DIRECT COSTS				**0**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**116,456**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **Equipment & Materials & Supplies (Rate: 16.0300, Base: 0) (Cont. on Comments Page)**					
TOTAL INDIRECT COSTS (F&A)				**83,544**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**200,000**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **200,000**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Burt Holzman

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET COMMENTS - Year 3

**** I- Indirect Costs**
Salaries (Rate: 77.4900, Base 106457)
Travel (Rate: 10.5000, Base 10000)

SUMMARY PROPOSAL BUDGET

Cumulative

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION: **Fermi Research Alliance, LLC**

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR: **Burt Holzman**

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Burt Holzman - Computing Serv Speciailist**	0.00	0.00	0.00	$ **0**	$
2. **Gabrielle Garzoglio - Applic Dev & Sys Analyst**	0.00	0.00	0.00	**0**	
3.					
4.					
5.					
6. ( ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**2**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.00	**0**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**3**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	34.58	0.00	0.00	**204,030**	
3. (**0**) GRADUATE STUDENTS				**0**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**204,030**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**115,338**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**319,368**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**30,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**0**	
TOTAL OTHER DIRECT COSTS				**0**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**349,368**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)					
TOTAL INDIRECT COSTS (F&A)				**250,632**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**600,000**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **600,000**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME: **Burt Holzman**

ORG. REP. NAME*: **Elizabeth Hisserich**

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

C ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

BUDGET JUSTIFICATION

Fermi National Accelerator Laboratory

Fermilab will be providing approximately 1.0 FTE of programming effort toward the proposed project titled "Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads". The compensation is consistent with similar work both within and outside of Fermilab. Personnel Cost estimates in Years 2 and 3 are calculated based upon a uniform escalation of 4.0%. Additional effort of 0.2 FTE will be provided in part by the glideinWMS project, funded by existing funds in CMS, in part by the Grid Services program, funded by the Fermilab Computing Division.

A. SENIOR PERSONNEL.

Burt Holzman will contribute project management effort in the context of the glideinWMS project. Dr. Holzman is a member of the CMS experiment and the current glideinWMS project lead. He is responsible for Grid Services support in US CMS as well as information services within the Open Science Grid.
Gabriele Garzoglio will act as a consultant in technical aspects of the project. Dr. Garzoglio is head of the Open Science Grid group in the Computing Division, leading the Grid Service program, and has years of experience in job, data, and information management on highly distributed systems for High Energy Physics.

B. OTHER PERSONNEL. Proposed compensation is consistent with that paid to other personnel engaged in similar work both within and outside Fermilab.

A junior programmer will work on the design, development, integration, and testing of the provisioning system, core of the grant proposal, and of the related monitoring infrastructure. He or she will be responsible for the end-to-end integration of the new system with existing Grid infrastructures, such as TeraGrid and Open Science Grid, focusing on the use cases of the CMS computing applications.

C. FRINGE BENEFITS
Benefits are requested at the rate of 56.53% of professional salaries. This includes vacation accrual rate (11%), OPTO (6.25%), and Fringe Benefits rate (33.5%).

D. PERMANENT EQUIPMENT
None

E. TRAVEL AND SUBSISTENCE.
Our travel budget of $10,000 per year covers domestic trips for collaboration and conferences.

G. OTHER DIRECT COSTS
None

I. TOTAL INDIRECT COSTS.
Fermilab's FY2009 provisional indirect cost rate is currently 77.49% (Salaries), 10.50% (Travel), and 16.03% (Other M&S) of MTDC, in accordance with Fermilab's contract with the Fermi Research Alliance, LLC (FRA) and the Department of Energy.

SUMMARY PROPOSAL BUDGET

YEAR 1

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of California-San Diego

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Frank Wuerthwein

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Frank Wuerthwein**	0.00	0.00	0.25	$ **2,478**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.25	**2,478**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	12.00	0.00	0.00	**8,130**	
3. (**1**) GRADUATE STUDENTS				**23,782**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**34,390**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**3,091**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**37,481**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**3,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**13,604**	
TOTAL OTHER DIRECT COSTS				**13,604**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**54,085**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **MTDC (Rate: 54.5000, Base: 40981)**					
TOTAL INDIRECT COSTS (F&A)				**22,335**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**76,420**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **76,420**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Frank Wuerthwein

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET

YEAR 2

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of California-San Diego

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Frank Wuerthwein

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Frank Wuerthwein - Associate Professor**	0.00	0.00	0.25	$ **2,654**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.25	**2,654**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	12.00	0.00	0.00	**8,178**	
3. (**1**) GRADUATE STUDENTS				**24,255**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**35,087**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**3,530**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**38,617**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**3,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**14,522**	
TOTAL OTHER DIRECT COSTS				**14,522**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**56,139**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **MTDC (Rate: 54.5000, Base: 42117)**					
TOTAL INDIRECT COSTS (F&A)				**22,954**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**79,093**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **79,093**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Frank Wuerthwein

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET

YEAR 3

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of California-San Diego

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Frank Wuerthwein

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Frank Wuerthwein**	0.00	0.00	0.25	$ **2,707**	$
2.					
3.					
4.					
5.					
6. (**0**) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.25	**2,707**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**1**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	12.00	0.00	0.00	**8,685**	
3. (**1**) GRADUATE STUDENTS				**26,806**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**38,198**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**3,624**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**41,822**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**3,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**15,503**	
TOTAL OTHER DIRECT COSTS				**15,503**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**60,325**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE) **MTDC (Rate: 54.5000, Base: 45321)**					
TOTAL INDIRECT COSTS (F&A)				**24,700**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**85,025**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **85,025**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Frank Wuerthwein

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

 ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

SUMMARY PROPOSAL BUDGET

Cumulative

FOR NSF USE ONLY		
PROPOSAL NO.	DURATION (months) Proposed	Granted
AWARD NO.		

ORGANIZATION
University of California-San Diego

PRINCIPAL INVESTIGATOR / PROJECT DIRECTOR
Frank Wuerthwein

A. SENIOR PERSONNEL: PI/PD, Co-PI's, Faculty and Other Senior Associates (List each separately with title, A.7. show number in brackets)	NSF Funded Person-months CAL	ACAD	SUMR	Funds Requested By proposer	Funds granted by NSF (if different)
1. **Frank Wuerthwein**	0.00	0.00	0.75	$ **7,839**	$
2.					
3.					
4.					
5.					
6. ( ) OTHERS (LIST INDIVIDUALLY ON BUDGET JUSTIFICATION PAGE)	0.00	0.00	0.00	**0**	
7. (**1**) TOTAL SENIOR PERSONNEL (1 - 6)	0.00	0.00	0.75	**7,839**	
B. OTHER PERSONNEL (SHOW NUMBERS IN BRACKETS)					
1. (**0**) POST DOCTORAL SCHOLARS	0.00	0.00	0.00	**0**	
2. (**3**) OTHER PROFESSIONALS (TECHNICIAN, PROGRAMMER, ETC.)	36.00	0.00	0.00	**24,993**	
3. (**3**) GRADUATE STUDENTS				**74,843**	
4. (**0**) UNDERGRADUATE STUDENTS				**0**	
5. (**0**) SECRETARIAL - CLERICAL (IF CHARGED DIRECTLY)				**0**	
6. (**0**) OTHER				**0**	
TOTAL SALARIES AND WAGES (A + B)				**107,675**	
C. FRINGE BENEFITS (IF CHARGED AS DIRECT COSTS)				**10,245**	
TOTAL SALARIES, WAGES AND FRINGE BENEFITS (A + B + C)				**117,920**	
D. EQUIPMENT (LIST ITEM AND DOLLAR AMOUNT FOR EACH ITEM EXCEEDING $5,000.)					
TOTAL EQUIPMENT				**0**	
E. TRAVEL 1. DOMESTIC (INCL. CANADA, MEXICO AND U.S. POSSESSIONS)				**9,000**	
2. FOREIGN				**0**	
F. PARTICIPANT SUPPORT COSTS					
1. STIPENDS $ **0**					
2. TRAVEL **0**					
3. SUBSISTENCE **0**					
4. OTHER **0**					
TOTAL NUMBER OF PARTICIPANTS (**0**) TOTAL PARTICIPANT COSTS				**0**	
G. OTHER DIRECT COSTS					
1. MATERIALS AND SUPPLIES				**0**	
2. PUBLICATION COSTS/DOCUMENTATION/DISSEMINATION				**0**	
3. CONSULTANT SERVICES				**0**	
4. COMPUTER SERVICES				**0**	
5. SUBAWARDS				**0**	
6. OTHER				**43,629**	
TOTAL OTHER DIRECT COSTS				**43,629**	
H. TOTAL DIRECT COSTS (A THROUGH G)				**170,549**	
I. INDIRECT COSTS (F&A)(SPECIFY RATE AND BASE)					
TOTAL INDIRECT COSTS (F&A)				**69,989**	
J. TOTAL DIRECT AND INDIRECT COSTS (H + I)				**240,538**	
K. RESIDUAL FUNDS				**0**	
L. AMOUNT OF THIS REQUEST (J) OR (J MINUS K)				$ **240,538**	$

M. COST SHARING PROPOSED LEVEL $ **0** AGREED LEVEL IF DIFFERENT $

PI/PD NAME
Frank Wuerthwein

ORG. REP. NAME*
Elizabeth Hisserich

FOR NSF USE ONLY		
INDIRECT COST RATE VERIFICATION		
Date Checked	Date Of Rate Sheet	Initials - ORG

C ***ELECTRONIC SIGNATURES REQUIRED FOR REVISED BUDGET**

Budget Justification, UCSD

The UCSD portion of the budget includes 10% of the annual Salary of Igor Sfiligoi, one graduate student, and 1 week of summer salary for the co-PI. In addition, it includes $3000 for travel, and $500 for miscellaneous expenses like supplies etc.

The travel funds will pay for travel to the twice a year project meetings for Sfiligoi, and the student, under the assumption that this implies one trip to FNAL, and a second trip to USC.

Igor Sfiligoi will be responsible for the day-to-day supervision of the graduate student. As Sfiligoi is part of the core glideinWMS development team, we include only a 10% incremental cost in this budget. The graduate student will be contributing to the program of work, especially in the area of automated error recovery, failure analysis, and testing. Würthwein's contribution lies in the high-level vision and direction. He is part of the computing management team of international as well as US-CMS, and his group at UCSD operates one of presently two large scale production glideinWMS systems worldwide. His contributions to the project are thus primarily in understanding the requirements and operational needs of large scale customers in general, and CMS in particular.

Current and Pending Support

(See GPG Section II.D.8 for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Ewa Deelman

Other agencies (including NSF) to which this proposal has been/will be submitted.

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

SDCI NMI Improvement: Pegasus: From Concept to Execution–Mapping Scientific Workflows onto the National Cyberinfrastructure

Source of Support: NSF

Total Award Amount: $1,700,00 Total Award Period Covered: 9/1/2007-8/31/2010

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 2.0 Acad: Sumr:

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Designing Scientific Software one Workflow at a Time

Source of Support: NSF

Total Award Amount: $313,556 Total Award Period Covered: 10/1/2007-9/30/2010

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 1.0 Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Center for Genomic Studies on Mental Disorders (CGSMD)

Source of Support: NIH

Total Award Amount: $2,250,000 Total Award Period Covered: 7/1/09-6/30/2012

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 1.0 Acad: Sumr:

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

CSR-AES: Collaborative Research: Intelligent Optimization of Parallel and Distributed Applications

Source of Support: NSF

Total Award Amount: 215,452 Total Award Period Covered: 8/1/2006-7/31/2009

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 0.4 Acad: Sumr:

Support: ☒ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

NHGRI EpiGenVar Coordinating Center.

Source of Support: NIH

Total Award Amount: $500,000 Total Award Period Covered: 9/1/2008-8/31/2012

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 1.0 Acad: Sumr:

---

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

Current and Pending Support

(See GPG Section II.D.8 for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Ewa Deelman

Other agencies (including NSF) to which this proposal has been/will be submitted.

Support: x Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:
The Genomic Psychiatry Cohort

Source of Support: NIH
Total Award Amount: $300,000 Total Award Period Covered: 10/1/2008-9/30/2013
Location of Project: USC/ISI
Person-Months Per Year Committed to the Project. Cal: 0.5 Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:
DC: Medium: Intelligent Data Placement in Support of Scientific Workflows

Source of Support: NSF
Total Award Amount: $969,133 Total Award Period Covered: 05/01/2009-04/30/2012
Location of Project: USC/ISI
Person-Months Per Year Committed to the Project. Cal: 0.96 Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title: CSR: Small: Meeting the Challenges of Large-Scale Computations on Heterogeneous Distributed

Source of Support: NSF
Total Award Amount: $489,292 Total Award Period Covered: 7/1/09-6/30/2012
Location of Project: USC/ISI
Person-Months Per Year Committed to the Project. Cal: 1.0 Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:
FutureGrid: An Experimental, High-Performance Grid Test-bed

Source of Support: NSF
Total Award Amount: $725,000 Total Award Period Covered: 10/1/2009-9/30/2013
Location of Project: USC/ISI
Person-Months Per Year Committed to the Project. Cal: 1.2 Acad: Sumr:

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:
STCI: Middleware for Monitoring and Troubleshooting of Large-Scale Applications on National Cyberinfrastructure

Source of Support: NSF
Total Award Amount: $1,875,830 Total Award Period Covered: 9/1/2009-8/31/2012
Location of Project: USC/ISI
Person-Months Per Year Committed to the Project. Cal: 1.2 Acad: Sumr:

---

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

Current and Pending Support

(See GPG Section II.D.8 for guidance on information to include on this form.)

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Ewa Deelman

Other agencies (including NSF) to which this proposal has been/will be submitted.

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Source of Support: NSF

Total Award Amount: $1,611,038 Total Award Period Covered: 9/1/09-8/31/2012

Location of Project: USC/ISI

Person-Months Per Year Committed to the Project. Cal: 1.2 Acad: Sumr:

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: Sumr:

---

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

Current and Pending Support

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Burt Holzman Pg 1 of 1

Other agencies to which this proposal has been/will be submitted.

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Source of Support: ISI - (Prime-National Science Foundation)

Total Award Amount: $200,000 Total Award Period Covered: 9/1/2009-8/31/2012

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 1.00 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

University of California, San Diego
Current and Pending Support

Frank Würthwein

(budgets include my fraction of the total of these grants)

Support Status: Current
Project Title: Elementary Particle Physics
Funding Source: DOE (co-PI of larger grant)
Total Award Amount: $250,000
Dates of Award: 1/1/2009-12/31/2009
Person-months of F.W. funded Per Year: 2 months

Support Status: Current
Project Title: DISUN
Funding Source: NSF (co-PI of larger grant)
Total Award Amount: $2,500,000
Dates of Award: 6/1/2005-5/31/2010
Person-months of F.W. funded Per Year: 0

Support Status: Current
Project Title: Open Science Grid
Funding Source: DOE (sub-contract via UW Madison)
Total Award Amount: $1,641,242
Dates of Award: 7/1/06 – 6/30/11
Person-months of F.W. funded Per Year: 0

Support Status: Current
Project Title: CMS Maintenance & Operations
Funding Source: NSF (sub-contract via UCLA)
Total Award Amount: $169,000
Dates of Award: 10/1/08 – 9/30/09
Person-months F.W. funded Per Year: 0

Support Status: Pending
Project Title: STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads
Funding Source: NSF (co-PI of larger grant)
Total Award Amount: $240,538
Dates of Award: 9/1/2009 – 8/31/2012
Person-months of F.W. funded Per Year: 0.25 months

Current and Pending Support

The following information should be provided for each investigator and other senior personnel. Failure to provide this information may delay consideration of this proposal.

Investigator: Gabriele Garzoglio Pg 1 of 1	Other agencies to which this proposal has been/will be submitted.

Support: ☐ Current ☒ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Source of Support: ISI - (Prime-National Science Foundation)

Total Award Amount: $200,000 Total Award Period Covered: 9/1/2009-8/31/2012

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.50 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: Acad: 0.00 Sumr: 0.00

Support: ☐ Current ☐ Pending ☐ Submission Planned in Near Future ☐ *Transfer of Support

Project/Proposal Title:

Source of Support:

Total Award Amount: $ Total Award Period Covered:

Location of Project:

Person-Months Per Year Committed to the Project. Cal: 0.00 Acad: 0.00 Sumr: 0.00

*If this project has previously been funded by another agency, please list and furnish information for immediately preceding funding period.

STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads

Facilities

Facilities at the University of Southern California Information Sciences Institute

USC/ISI is a large research center with an emphasis on projects that blend basic and applied research through exploratory system development. Many ISI researchers, including the PI, are also research faculty in the Department of Computer Science, and advise graduate students on research projects and dissertations. More than sixty graduate students at USC complete their degrees as research assistants on ISI projects. Many undergraduates do internships at ISI.

USC's High Performance Computing Facility currently is currently ranked 24th on the Top 500 computer ranking and second fastest in academia, with over 10 Teraflops of computing, 1,830 nodes, and over 30 Terabytes of storage. In addition to the campus wide resource, ISI operates a local Condor pool of approximately 50 nodes, and has exclusive access to a 2.5 TB, 16 node storage cluster, a 6 TB parallel network attached storage facility and a 9 processor computational/visualization cluster. At ISI, an institute-wide 90 node dual-processor general-purpose compute cluster was installed. Equipment procured for exclusive use in the proposed project will build on this computational framework, leveraging institute investment in management, network, and power. These facilities are directly connected to the ISI backbone network with dedicated multiple Gb/sec network connections. Additionally, the PIs at ISI have their own compute resources which include a heterogeneous cluster with 6 dual-core Linux-based system and a MS Windows installation as well as an 9-node dual-cpu linux-based cluster.

Facilities at Fermilab

Fermilab's Scientific Computing Facilities provide the broad Fermilab user community (including accelerator and accelerator R&D scientists) with state-of-the-art large scale computing, storage, networking and database resources to meet their needs. The computing facilities include.

The following facilities are relevant to the work of this proposal and interfaces can be provided as needed

- The LHC CMS experiment's Tier-1 in the US which archives data for later distribution to the several hundred CMS physicists in the US and several thousand members worldwide, as well as provides significant data reprocessing and analysis capabilities.
- The Fermilab Campus Grid, FermiGrid, which provides access to high throughput computing farms which serve the data processing needs of all the users, and interface to the multi-disciplinary national grid infrastructure, the Open Science Grid.
- Tape storage facilities for use by all Fermilab users and large disk farms as tape front-ends and online data caching systems.

The current total capacity of the Fermilab facilities is 30,000 processing cores, 13 Petabytes of archival tape storage, easily expandable to 40 Petabytes, and more than 3.5 Petabytes of disk storage. The processors are used at 85% of capacity; Typical tape writing is at 100 Terabytes a day and reading is at >100 Terabytes a day. Data transfers are on average >2Gbit/sec to and from the Wide Area Network (with frequent bursts of >10Gb/sec). The monthly total of data transferred between Fermilab and the WAN is thus >0.65 Petabyes per month inbound and 2.2 Petabyes per month outbound.

Facilities at the University of California High Energy Physics Group

The High Energy Physics group at UCSD has a strong emphasis on software and computing. We have 12 people working in this area on topics including hardware operations, distributed computing middleware development, integration, and operation, application development in

particle physics, and application support of a wide range of domain sciences in the context of the Open Science Grid.

Our computing infrastructure is an APC built hot-isle enclosure that we are upgrading this summer to support 12 racks of computing hardware. The presently deployed 5 racks include 1000 cores of batch computing and 300TB of disk space, both accessed via OSG APIs, 5 Dual Quad core login nodes to support software development and interactive analysis by the CMS groups from UCSB, UCSD, and UCR, and about 1 rack supporting infrastructure like the OSG Compute and Storage Element middleware servers, a Squid server, hadoop file system namespace server, CMS data management infrastructure, a 5 node glideinWMS server infrastructure to support globally distributed analysis for CMS, entry points for our scalability testing overlay testbed of 5000 batch slots, Xen nodes to support virtualization, etc. The networking infrastructure is provided by a Cisco 6509 with 10Gbps WAN uplinks to ESNet and CENIC each. Traffic to FNAL is routed via ESNet directly, whereas all other traffic is routed via CENIC. The two networks are configured for automatic failover. Our storage infrastructure is shown to support sinking in excess of 10Gbps via the WAN during normal production operations. The dominant use case of our storage is as cache for CMS data analysis.

For the work described in this proposal, we expect to provide a development platform, experience from the present glideinWMS production platform, access to the scalability testing infrastructure, access to the main production cluster, and of course desktop computing for our personnel. If needed, we could provide a small Xen platform for development of the cloud computing capabilities of CorralWMS.

Southern California Earthquake Center
an NSF+USGS center
University of Southern California, 3651 Trousdale Parkway, Suite 169, Los Angeles, CA 90089-0742
Phone: 213-740-5843 Fax: 213-740-0011 E-Mail: scec@usc.edu Web: www.scec.org

May 26, 2009

Dr. Ewa Deelman
Information Sciences Institute
University of Southern California
4676 Admiralty Way, Suite 1001
Marina Del Rey, CA 90092

Dear Dr. Deelman:

We would like to confirm our continued commitment to collaborating with you in the area of resource provisioning in distributed systems such as the National Cyberinfrastructure. We fully support the work described in your NSF proposal "STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads". Today we use your Corral resource provisioning service to execute our CyberShake workflows on the TeraGrid.

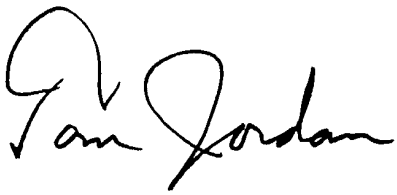
As you know, CyberShake is one of the most ambitious computational programs in geophysics. Its goal is to replace empirical estimates of seismic motions from large earthquakes with physics-based calculations that take into account geologic structure and rupture complexity. If successful, CyberShake will transform how probabilistic seismic hazard analysis is formulated and applied to reducing earthquake risk. Our preliminary calculations show that CyberShake is computational feasible and scientifically exciting.

Our CyberShake workflows are very complex and composed of approximately 800,000 individual tasks. Although we use the Pegasus Workflow Management System to restructure the workflow and reduce its complexity, we still face the problem of using resources such as the TeraGrid efficiently. These resources are geared primarily towards long duration, multi-core programs rather than large numbers of single-core computations. However, these resources are critical to providing a reasonable time to solution for our application. One approach we have taken to leverage the high-performance TeraGrid resources is to provision resources ahead of the workflow execution. In this way the execution system is used efficiently by allocating a number of cores for the duration of the workflow and at the same time CyberShake can run efficiently and make full use of the resources. One could potentially provision the desired resources manually (as we have done in the past), but this is a very labor intensive and error-prone process.

Your Corral service which automates the resources provisioning and "watches" over the provisioned resources is thus critical to providing us with an efficient way of running our workflows on the TeraGrid. The proposed integration of Corral with glideinWMS developed at Fermi Lab can potentially also enable us to run our computations on the Open Science Grid and thus make more resources available to our earthquake system science applications. As a result we can also leverage the proposed technologies to enable more SCEC applications to run across the National Cyberinfrastructure.

We wish you the very best of luck with your proposal and look forward to our continued collaboration.

Sincerely,



Thomas H. Jordan
Director, Southern California Earthquake Center
University Professor and W. M. Keck Foundation Professor of Earth Sciences

Renci
Renaissance Computing Institute
Catalyst for Innovation

May 25, 2009

Dr. Ewa Deelman
Information Sciences Institute
University of Southern California
4676 Admiralty Way, Suite 1001
Marina Del Rey, CA 90092

Dear Dr. Deelman:

We would like to confirm our continued commitment to collaborating with you in the area of resource provisioning in distributed systems within the RENCI Engagement Team. As such we fully support the work described in your NSF proposal "STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads". We believe that a comprehensive and robust resource provisioning system can make it easier for us to deliver efficient usage of the resources provided by OSG and the TeraGrid. Today, CMS is one of the main users of glideinWMS on OSG but there are also other applications that we are working with as part of the OSG Engagement program that could benefit from the enhanced capabilities of glideinWMS and its integration with Corral. Additionally, the RENCI TeraGrid Science Gateway will likely benefit from this work as well.

The Engagement Team works with users from many science domains to leverage OSG. During the previous year, we have worked with a diverse group of more than 20 scientists and teams, delivering more than 2.5 million opportunistic OSG cycles. Services such as the proposed enhanced glideinWMS enable us to deliver better solutions more quickly for researchers that rely upon computational models.

We are also excited about the capabilities of the proposed system to bridge the national cyberinfrastructure as well as cloud technologies, and we are looking forward to our users being able to make use of TeraGrid resources as appropriate.

We wish you success with your proposal and look forward to our continued collaboration.

Best regards,

John McGee
OSG Engagement Coordinator
Manager, Cyberinfrastructure Development
Renaissance Computing Institute
mcgee@renci.org

Computer Sciences Department

1210 West Dayton St.
Madison, WI 53706

Miron Livny
Phone - (608) 262-0856
Fax - (608) 262-9777
Email - miron@cs.wisc.edu

May 24, 2009

Dr. Ewa Deelman
USC Information Sciences Institute
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90092

Dear Ewa,

This letter is to express my commitment to sustain the support the Condor project at University of Wisconsin-Madison has been providing to your group and the other groups who are part of the program of work of the "STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads" proposal you are leading. We regard the use of glideinWMS by High Energy Physics (CMS, CDF, MINOS) and the use of Corral by SCEC and Pegasus WMS by LIGO as import and integral parts of the Condor community that we are committed to support. We are looking forward to working with your team on the integration of Coral and glideinWMS and are committed to do our best to enhance our technologies to meet the needs of the science communities that have and will adopt your integrated tools.

Best wishes for the success of the proposal.

Sincerely,

Professor Miron Livny
Director of the UW Center for High Throughput Computing
Computer Sciences Department
University of Wisconsin-Madison

Fermilab

Ian Fisk
CD
CMS
630.840.6467

ifisk@fnal.gov

May 22, 2009

Dr. Ewa Deelman
Information Sciences Institute
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90292

Dr. Burt Holzman
Fermilab
PO Box 500
Batavia, IL 60510

Dear Ewa and Burt,

This is a letter in strong support of the proposal entitled "Integrated Resource Provisioning Across National Cyberinfrastructure in Support of Scientific Workloads." submitted to the NSF STCI Program. I believe the proposed work would be immediately beneficial to my field and to a growing number of organizations.

CMS has already significantly benefited from one component of your proposal, the GlideinWMS, over the last two years. In CMS the scale of processing has increased by roughly a factor of two each year during the final 3 years of preparation for the start of the experiment, and we expect this trend to continue for at least the first several years of running. While the GlideinWMS has been instrumental in achieving the large scale organized processing we currently enjoy, we see that there is clearly continued development work needed to reach the challenging new scales and to support new use-cases. I look forward to seeing the results of the strong team you are assembling.

The combination of the dynamic provisioning available in GlideinWMS and the proactive provisioning performed by Corral seems like a natural fit. I expect that this combination will be a great benefit to communities in the Open Science Grid, including CMS, as well as communities in the TeraGrid with large-scale processing needs.

Sincerely,

Ian Fisk

Deputy US-CMS Software and Computing Manager

Open Science Grid

May 28, 2009

Dear Ewa,

I am writing as Executive Director of the Open Science Grid, on behalf of the OSG Consortium, to express my strong support for the Information Sciences Institute's "STCI: Integrated Resource Provisioning Across the National Cyberinfrastructure in Support of Scientific Workloads" proposal submitted to the NSF STCI program.

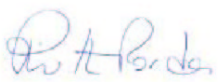
The Open Science Grid Consortium's mission is to provide a national distributed computing infrastructure that is open to all of science and in particular is crucially relied on by the physics collaborations in the US. US ATLAS, US CMS, LIGO, D0, CDF and STAR, depend on and contribute to the work of the OSG. Virtual Organizations in several other science areas, including biology, weather research and nanotechnology, are expanding its' use into the broader research arenas. One of the current impediments to widespread use of OSG by other science disciplines is the ability to easily set up and run complex, application domain job sets. Your proposed program of work directly addresses these needs.

As the complexity of the applications using the OSG and the national cyberinfrastructure (CI), and as the scale of the infrastructure increases, there are crucial needs for systems that will provide scientists across many different disciplines ways to easily specify, run, and manage structured job execution across all technologies in the CI. Your Integrated Resource Provisioning proposed program of work provides an evolutionary and integrating approach to existing solutions, providing a common, supported solution to these needs.

We understand that you plan close collaboration with the OSG, including use of existing OSG elements and workflow management systems, to develop the improved resource provisioning system. We look forward to assisting in providing requirements, review of the design and development, performance testing and deployment of the resulting service. We also plan to integrate the necessary components into the OSG software stack, through the Virtual Data Toolkit, to make the service available to additionally existing and new OSG science communities. Training on the new system will be added to our Grid School curriculum.

I am confident that the close involvement of the principals on the Integrated Resource Provisioning proposal with the activities of the OSG Consortium means that successful collaboration is virtually guaranteed. Additionally, that the needs of the OSG stakeholders will be addressed and developed by the Integrated Resource Provisioning program of work as it proceeds.

With Regards,



Ruth Pordes
Executive Director

On behalf of the OSG Consortium